

**SEARCH REQUEST FORM**

Scientific and Technical Information Center

Requester's Full Name: \_\_\_\_\_ Examiner #: \_\_\_\_\_ Date: \_\_\_\_\_  
Art Unit: \_\_\_\_\_ Phone Number 30 \_\_\_\_\_ Serial Number: \_\_\_\_\_  
Mail Box and Bldg/Room Location: \_\_\_\_\_ Results Format Preferred (circle): PAPER DISK E-MAIL

**If more than one search is submitted, please prioritize searches in order of need.**

\*\*\*\*\*

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: \_\_\_\_\_

Inventors (please provide full names): \_\_\_\_\_

Earliest Priority Filing Date: \_\_\_\_\_

*\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

Point of Contact:  
Jan Delaval  
Librarian-Physical Sciences  
CM1 1E04 Tel: 308-4498

\*\*\*\*\*

**STAFF USE ONLY**

	Type of Search	Vendors and cost where applicable
Searcher: <u>Jan</u>	NA Sequence (#) <u>6</u>	STN _____
Searcher Phone #: <u>4498</u>	AA Sequence (#) <u>6</u>	Dialog _____
Searcher Location: _____	Structure (#) _____	Questel/Orbit _____
Date Searcher Picked Up: <u>11/7</u>	Bibliographic _____	Dr. Link _____
Date Completed: <u>11/9</u>	Litigation _____	Lexis/Nexis _____
Searcher Prep & Review Time: _____	Fulltext _____	Sequence Systems <u>✓</u>
Clerical Prep Time: <u>20</u>	Patent Family _____	WWW/Internet _____
Online Time: <u>10</u>	Other _____	Other (specify) _____



Searches for User *mjamroz* (Count = 440)

Queries 391 through 440.

Latest

Prev

Next

Oldest

Edit

Help

Return

Main Menu

Logout

09/06/08, 02:11

S #	Updt	Database	Query	Time	Comment
<u>S440</u>	<u>U</u>	USPT,PGPB	(endoglin same antibod\$ ) same monoclonal	2001-11-09 15:39:06	
<u>S439</u>	<u>U</u>	USPT,PGPB	endoglin same antibod\$	2001-11-09 15:38:39	
<u>S438</u>	<u>U</u>	USPT,PGPB	((TGF) adj beta adj binding adj protein ) same antibod\$	2001-11-09 15:33:19	
<u>S437</u>	<u>U</u>	USPT,PGPB	(TGF) adj beta adj binding adj protein	2001-11-09 15:32:52	
<u>S436</u>	<u>U</u>	USPT	5589375.pn.	2001-11-09 15:31:35	
<u>S435</u>	<u>U</u>	USPT	((protein adj tyrosine adj phosphatase) same antibod\$ ) and @py<1997 ) and claim	2001-11-09 12:58:12	
<u>S434</u>	<u>U</u>	USPT	((protein adj tyrosine adj phosphatase) same antibod\$ ) and @py<1999	2001-11-09 12:56:42	
<u>S433</u>	<u>U</u>	USPT	((protein adj tyrosine adj phosphatase) same antibod\$ ) and @py<1997	2001-11-09 12:24:41	
<u>S432</u>	<u>U</u>	USPT	(protein adj tyrosine adj phosphatase) same antibod\$	2001-11-09 12:24:12	
<u>S431</u>	<u>U</u>	USPT	5580742.pn.	2001-11-09 12:07:49	
<u>S430</u>	<u>U</u>	USPT	(antibody same phosphotyrosine same phosphatase ) and @py<1997	2001-11-09 11:59:41	
<u>S429</u>	<u>U</u>	USPT	antibody same phosphotyrosine same phosphatase	2001-11-09 11:59:24	
<u>S428</u>	<u>U</u>	USPT	(antibody same phosphotyrosine same (kinase or phosphatase) ) and @py<1997	2001-11-09 11:58:53	
<u>S427</u>	<u>U</u>	USPT	(antibody same phosphotyrosine same (kinase or phosphatase) ) same @py<1997	2001-11-09 11:58:42	
<u>S426</u>	<u>U</u>	USPT	antibody same phosphotyrosine same (kinase or phosphatase)	2001-11-09 11:57:52	
<u>S425</u>	<u>U</u>	USPT	5863781.pn.	2001-11-09 10:22:54	





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 8, 2001, 19:25:30 ; Search time 438.73 Seconds

(without alignments)  
3293.145 Million cell updates/sec

Title: US-09-668-021-1

Perfect score: 2301  
Sequence: 1 agagccctgctactggaag.....caatgaatcatgaccgaag 2301

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Archived: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

N\_Geneseq\_0601.\*  
1: /SIDSI/gcgdata/geneseq/geneseq/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq/NA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
1	2301	100.0	2301	AAA29055	Human TGF-beta bin
2	2299.4	99.9	2301	AAA29061	Mutant human TGF-b
3	2299.4	99.9	2301	AAA29062	Human TGF-beta bin
4	2297.8	99.9	2301	AAA29056	Human TGF-beta bin
5	2280.6	99.1	2329	AAA1073	Human secreted pro
6	2047.8	89.0	2329	AAA9409	Human DAN/Cerberus
7	2047.8	89.0	9301	AAA29054	Human TGF-beta bin
8	642	27.9	642	AAA94051	Human DAN/Cerberus
9	613.2	26.6	642	AAA29057	Vervet TGF-beta bl
10	536.2	23.3	674	AAA29059	Rat TGF-beta blndi
11	502	21.8	638	AAA29058	Murine TGF-beta bl

12	467.8	20.3	532	21	AAA29060	Bovine TGF-beta bl
13	442	18.3	1104	21	AAA94050	Human DAN/Cerberus
14	372.2	16.2	35828	21	AAA29063	Murine TGF-beta bl
15	193.8	8.4	198	20	AAA40842	Human secreted pro
16	193.4	8.4	196	21	AA05741	Oligonucleotide D1
17	100.4	4.4	936	22	AAE58252	Oligonucleotide D1
18	100.4	4.4	936	22	AAE58254	Oligonucleotide D1
19	100.4	4.4	936	22	AAE58257	Oligonucleotide D1
20	100.4	4.4	936	22	AAE58259	Oligonucleotide D2
21	100.4	4.4	936	22	AAE58262	Oligonucleotide D2
22	100.4	4.4	936	22	AAE58255	Oligonucleotide D1
23	98.6	4.3	936	22	AAE58252	Oligonucleotide D1
24	98.6	4.3	936	22	AAE58254	Oligonucleotide D1
25	98.6	4.3	936	22	AAE58257	Oligonucleotide D1
26	98.6	4.3	936	22	AAE58259	Oligonucleotide D2
27	98.6	4.3	936	22	AAE58262	Oligonucleotide D2
28	98.6	4.3	936	22	AAE58255	Oligonucleotide D1
29	82.6	3.6	954	22	AAE58257	Oligonucleotide D1
30	82.6	3.6	962	21	AAE58255	Oligonucleotide D1
31	82.6	3.6	962	21	AAE58257	Oligonucleotide D1
32	82.6	3.6	962	22	AAE58259	Oligonucleotide D2
33	82.6	3.6	962	22	AAE58262	Oligonucleotide D2
34	69.6	3.0	900	18	AAE47661	Human cell CDNA, SE
35	69.6	3.0	900	19	AAE77255	Human cell CDNA, SE
36	69.6	3.0	1692	21	AAE79467	Human small CNV-11
37	69.6	3.0	1710	20	AAE84238	CDNA sequence of h
38	69.6	3.0	1719	21	AAE84238	Human adult retina
39	69.6	3.0	1774	20	AAE07571	Human adult retina
40	66.4	2.9	114955	20	AAE53491	Human adenovirus A1
41	57.8	2.5	412	22	AAE66218	Novel human polyom
42	57.4	2.5	244	22	AAE58238	Oligonucleotide D1
43	55.6	2.4	114955	20	AAE53491	Human adenovirus A1
44	53	2.3	244	22	AAE58238	Oligonucleotide D1
45	52.4	2.3	2943	17	AAE16480	S88 virus gp glyco

#### ALIGNMENTS

RESULT 1	
AAA29055	
ID AAA29055 strand; CDNA: 2301 BP.	
AC AAA29055;	
DT 12-SEP-2000 (first entry)	
XX	
DE Human TGF-beta binding protein (BEER) CDNA.	
XX	
KM osteopathic: transforming growth factor-beta; TGF-beta: binding protein;	
KM BEER: chromosome 17q12-21; gene therapy; antisense therapy; fracture;	
KW bone mineralization; ss.	
XX	
OS Homo sapiens.	
XX	
FT Key	Location/Qualifiers
FT CDS	48..689
FT	/tag= a
FT	/product= TGF-beta_binding_protein
XX	
PN MO200032773-A1.	
XX	
PD 08-JUN-2000.	
XX	
PE 24-NOV-1999;	99WO-US27990.
XX	
PR 27-NOV-1998;	98US-0110283.
XX	
PA (DARW-) DARWIN DISCOVERY LTD.	
XX	
PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;	
PI Van Ness J, Winkler DG;	
XX	

*OKM 5*

DR WP1: 2000-412321/35.  
DR P-PSDB: AAY6429.

PT Nucleic acids (1) encoding a transforming growth factor beta binding  
protein, useful for identifying agents for treating osteopenia,  
osteoporosis and fractures

PS Claim 1: Page 114-115: 162pp: English.

XX This cDNA encodes a human transforming growth factor-beta (TGF-beta)  
CC binding protein designated BEER. The hBEER gene has been localized  
CC to the chromosome 17q12-21. The cDNA and protein may be used for  
CC prevention, treatment and diagnosis of diseases associated with  
CC inappropriate BEER expression. For example, they may be used to treat  
CC disorders associated with decreased TGF-beta BP expression. The cDNA or  
CC vectors may be administered to treat diseases by rectifying mutations or  
CC deletions in a patient's genome that affect the activity of BEER by  
CC expressing inactive proteins or to supplement the patient's own production  
CC of BEER polypeptides. The nucleic acids may be used for recombinant  
CC production of BEER, gene therapy, antisense therapy, as probes for  
CC diagnostic assays and for functional studies. BEER may be used to raise  
CC antibodies and for identification of BEER modulators. BEER antagonists  
CC may be used to increase bone mineral content for the treatment of  
CC disorders such as osteopenia, osteoporosis, fractures and other  
disorders associated with low mineral content.

XX Sequence 2301 BP; 568 A; 635 C; 614 G; 484 T; 0 other:

Query Match 100.0%; Score 2301; DB 21; Length 2301;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 agagcctggtgacacggaagtggtggtggtcctcctcgtggtggtacatgacatccac 60  
DB 1 agagcctggtgacacggaagtggtggtggtcctcctcgtggtggtacatgacatccac 60  
QY 61 tggcctgtgtctcgtctgtcgtgtgtacacagccttcgtgtgtgtgtgtgtgtgtgtgt 120  
DB 61 tggcctgtgtctcgtctgtcgtgtgtgtacacagccttcgtgtgtgtgtgtgtgtgtgtgt 120  
QY 121 ggtggaagcgttcaagaatgtgtcagcaagatcatcccgagctcgtggaatctacccg 180  
DB 121 ggtggaagcgttcaagaatgtgtcagcaagatcatcccgagctcgtggaatctacccg 180  
QY 121 ggtggaagcgttcaagaatgtgtcagcaagatcatcccgagctcgtggaatctacccg 180  
DB 121 ggtggaagcgttcaagaatgtgtcagcaagatcatcccgagctcgtggaatctacccg 180  
QY 181 agcctccacggagctgtggaacaagacatgaaacgggagaaacgggagcg 240  
DB 181 agcctccacggagctgtggaacaagacatgaaacgggagaaacgggagcg 240  
QY 181 agcctccacggagctgtggaacaagacatgaaacgggagaaacgggagcg 240  
DB 181 agcctccacggagctgtggaacaagacatgaaacgggagaaacgggagcg 240  
QY 241 ctccacacaccccttgaagacaagacgtgtccagatcagctgctcgcaagcttgcact 300  
DB 241 ctccacacaccccttgaagacaagacgtgtccagatcagctgctcgcaagcttgcact 300  
QY 241 ctccacacaccccttgaagacaagacgtgtccagatcagctgctcgcaagcttgcact 300  
DB 241 ctccacacaccccttgaagacaagacgtgtccagatcagctgctcgcaagcttgcact 300  
QY 301 tcaacccgacgttgaacgagtggtggtggtggtggtggtggtggtggtggtggtggt 360  
DB 301 tcaacccgacgttgaacgagtggtggtggtggtggtggtggtggtggtggtggtggt 360  
QY 301 tcaacccgacgttgaacgagtggtggtggtggtggtggtggtggtggtggtggtggt 360  
DB 301 tcaacccgacgttgaacgagtggtggtggtggtggtggtggtggtggtggtggtggt 360  
QY 361 gctcggcgacgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 420  
DB 361 gctcggcgacgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 420  
QY 421 ggtggaagcgttgaacgagtggtggtggtggtggtggtggtggtggtggtggtggt 480  
DB 421 ggtggaagcgttgaacgagtggtggtggtggtggtggtggtggtggtggtggtggt 480  
QY 421 ggtggaagcgttgaacgagtggtggtggtggtggtggtggtggtggtggtggtggt 480  
DB 421 ggtggaagcgttgaacgagtggtggtggtggtggtggtggtggtggtggtggtggt 480  
QY 481 agctgtgtgtccggt 540  
DB 481 agctgtgtgtccggt 540  
QY 541 gcaagtgcaagcgttcaacccgttccacacacagtcgtggaagtcgaagtcgtggaacg 600  
DB 541 gcaagtgcaagcgttcaacccgttccacacacagtcgtggaagtcgaagtcgtggaacg 600

QY 601 agcgcctcgtgacgacagaaggtccggaagcgtggtggtggtggtggtggtggtggtggt 660  
DB 601 agcgcctcgtgacgacagaaggtccggaagcgtggtggtggtggtggtggtggtggtggt 660  
QY 661 accagcgtgacgtggaagacgttgaagcgtggtggtggtggtggtggtggtggtggtggt 720  
DB 661 accagcgtgacgtggaagacgttgaagcgtggtggtggtggtggtggtggtggtggtggt 720  
QY 721 gcccggcccttgaacccggtggtggtggtggtggtggtggtggtggtggtggtggtggt 780  
DB 721 gcccggcccttgaacccggtggtggtggtggtggtggtggtggtggtggtggtggtggt 780  
QY 781 attcatgttaagtctgtgcaacacgaggtggtggtggtggtggtggtggtggtggtggtggt 840  
DB 781 attcatgttaagtctgtgcaacacgaggtggtggtggtggtggtggtggtggtggtggtggt 840  
QY 841 atcccggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 900  
DB 841 atcccggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 900  
QY 901 gagggaattgagatcagac 960  
DB 901 gagggaattgagatcagac 960  
QY 961 tgggtgttcccttccac 1020  
DB 961 tgggtgttcccttccac 1020  
QY 1021 agggagcgtgtgtggaagtggtggaagtggtggaagtggtggaagtggtggaagtggtgga 1080  
DB 1021 agggagcgtgtgtggaagtggtggaagtggtggaagtggtggaagtggtggaagtggtgga 1080  
QY 1081 ccccttgcac 1140  
DB 1081 ccccttgcac 1140  
QY 1141 caactgtgagatgt 1200  
DB 1141 caactgtgagatgt 1200  
QY 1201 tacaacatctctcctcgtggaac 1260  
DB 1201 tacaacatctctcctcgtggaac 1260  
QY 1261 taggaatcgtggaagtggtggaagtggtggaagtggtggaagtggtggaagtggtggaag 1320  
DB 1261 taggaatcgtggaagtggtggaagtggtggaagtggtggaagtggtggaagtggtggaag 1320  
QY 1321 cagaagtggaagtggtggaagtggtggaagtggtggaagtggtggaagtggtggaagtggt 1380  
DB 1321 cagaagtggaagtggtggaagtggtggaagtggtggaagtggtggaagtggtggaagtggt 1380  
QY 1381 caaggtcacttccagaaatcagaaatcagaaatcagaaatcagaaatcagaaatcagaaatc 1440  
DB 1381 caaggtcacttccagaaatcagaaatcagaaatcagaaatcagaaatcagaaatcagaaatc 1440  
QY 1441 caaagcaaaaaaaagtggaagtggtggaagtggtggaagtggtggaagtggtggaagtggt 1500  
DB 1441 caaagcaaaaaaaagtggaagtggtggaagtggtggaagtggtggaagtggtggaagtggt 1500  
QY 1501 tccctggaagtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 1560  
DB 1501 tccctggaagtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 1560  
QY 1561 cccctcattcctcaaaatcattcattcattcattcattcattcattcattcattcattcattc 1620  
DB 1561 cccctcattcctcaaaatcattcattcattcattcattcattcattcattcattcattcattc 1620  
QY 1621 ggtggaagtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 1680  
DB 1621 ggtggaagtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 1680  
QY 1681 acccatagcattgtttaaagtac 1740

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 9, 2001, 00:34:26 : Search time 222.63 Seconds  
(Without alignments)  
2340.771 Million cell updates/sec

Title: US-09-668-021-5  
Perfect score: 2301  
Sequence: 1 agagcctgtctactctgaag.....caatgaatcatcagccgaag 2301

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Archived: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PTCUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	82.6	3.6	962	3	US-09-188-930-40
2	82.6	3.6	962	3	US-09-188-930-209
3	69.6	3.0	900	1	US-08-468-847B-1
4	52.4	2.3	2943	1	US-08-062-747A-7
5	49.6	2.2	4257	2	US-08-690-473-1
6	49.6	2.2	4257	4	US-09-259-821A-1
7	49.6	2.2	4257	4	US-08-843-659-1
8	49.6	2.2	12001	1	US-08-458-568A-11
9	48.8	2.1	71989	4	US-09-443-501A-2
10	48.8	2.1	7218	1	US-08-232-463-14
11	47.6	2.1	2150	2	US-08-318-837-1
12	46.8	2.0	1910	4	US-09-593-711A-3
13	46.8	2.0	1914	1	US-07-601-094-1
14	46.8	2.0	1914	1	US-08-012-735-1
15	46.6	2.0	4403765	4	US-09-103-840A-2
16	44.2	1.9	33529	4	US-09-144-085-3
17	43.4	1.9	8438	1	US-07-945-283-1
18	43.4	1.9	34303	2	US-08-735-609-4
19	43.4	1.9	34303	2	US-08-735-609-4
20	43.4	1.9	34303	3	US-09-315-972-4
21	43.4	1.9	34303	3	US-09-244-752-4
22	43.4	1.9	34303	3	US-09-245-497-4
23	43.4	1.9	34382	2	US-08-374-483-6
24	43.4	1.9	35408	4	US-08-973-334-3
25	43.4	1.9	35408	4	US-09-563-869A-3
26	43.4	1.9	35408	4	US-08-549-489-3
27	43.4	1.9	35935	2	US-08-735-609-1

c	28	43.4	1.9	35935	2	US-08-735-609-1	Sequence 1, Appl
c	29	43.4	1.9	35935	3	US-08-379-452-43	Sequence 43, Appl
c	30	43.4	1.9	35935	3	US-09-315-972-1	Sequence 1, Appl
c	31	43.4	1.9	35935	3	US-09-244-752-1	Sequence 1, Appl
c	32	43.4	1.9	35935	3	US-09-245-497-1	Sequence 1, Appl
c	33	43.4	1.9	35935	3	US-09-409-670-43	Sequence 43, Appl
c	34	43.2	1.9	1578	1	US-08-681-129-1	Sequence 1, Appl
c	35	42.8	1.9	2634	1	US-08-196-218-31	Sequence 31, Appl
c	36	42.8	1.9	2634	1	US-08-681-953-31	Sequence 31, Appl
c	37	42.8	1.9	13842	4	US-09-105-537-30	Sequence 30, Appl
c	38	42.8	1.9	36778	4	US-09-320-878-19	Sequence 19, Appl
c	39	42.8	1.9	38506	3	US-08-417-822A-1	Sequence 1, Appl
c	40	42.6	1.9	835	1	US-08-396-218-1	Sequence 1, Appl
c	41	42.6	1.9	1269	1	US-08-760-116-1	Sequence 1, Appl
c	42	42.6	1.9	1269	1	US-07-642-734C-3	Sequence 3, Appl
c	43	42.6	1.9	20235	3	US-08-439-009A-3	Sequence 3, Appl
c	44	42.6	1.9	20235	3	US-08-242-677-1	Sequence 1, Appl
c	45	42.2	1.8	5173	1	US-08-242-677-1	Sequence 1, Appl

#### ALIGNMENTS

RESULT 1  
US-09-188-930-40  
Sequence 40, Application US/09188930A  
Patent No. 6150502  
GENERAL INFORMATION:  
APPLICANT: Watson, James D.  
APPLICANT: Strachan, Lorna  
APPLICANT: Sleeman, Mathew  
APPLICANT: Olmstead, Rene  
APPLICANT: Murison, James Greg  
TITLE OF INVENTION: Compositions Isolated From Skin Cells  
TITLE OF INVENTION: and Methods for Their Use  
FILE REFERENCE: 11000.1011c1  
CURRENT FILING DATE: 1998-11-09  
NUMBER OF SEQ ID NOS: 348  
SOFTWARE: FASTSEQ for Windows Version 3.0  
SEQ ID NO 40  
LENGTH: 962  
TYPE: DNA  
ORGANISM: mouse  
US-09-188-930-40

Query Match 3.6%: Score 82.6; DB 3: Length 962;  
Best Local Similarity 52.5%: Pred. NO. 1.3e-11;  
Matches 287; Conservative 0; Mismatches 239; Indels 21; Gaps 4;

QY	123	tgccagcgcttcaagatgacgacgaatcaccgagacgtcgagagtaaccccgag	182
DB	107	tggttggttttaaaatgacgacgaatccttattcaatggtgtaacctgic	166
QY	183	ccttcacgagcggtcgaggaacacacacatgacgagcgaggaagagcgagcct	242
DB	167	ccggacac---accacgacgacacacacacccctgtaacgacgagtaggacgagc	223
QY	243	ccccccaccc---cttgaagaccacgaagcgtgacgagtaacgctccgagcgac	299
DB	224	ttcagtagcactgtagctgacgacgacacacagtcagtgagtgccgaggaacg	283
QY	300	ttcaccgcttgcgtgacgacgagcggtgacgagcgacacggttaccgagcggtg	359
DB	284	ttcaccacaaatcattcgtgacgagcgacgacgacgacgacgacgacgagcggtg	343
QY	360	tgcttcgagcgagtgagcgagcgagcgagcgagcgagcgagcgagcgagcgagc	413
DB	344	tgccgagcgagtgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagc	403
QY	414	-----aagtgtagcgacacacgagcgagcgagcgagcgagcgagcgagcgagc	467



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 8, 2001, 20:59:11 ; Search time 5549.44 Seconds  
(without alignments)  
3919.497 Million cell updates/sec

Title: US-09-668-021-5  
Perfect score: 2301  
Sequence: 1 agagcctgtctactgtgaag.....caatgaatcatgaccgaaag 2301

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Archived: 1022815 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*  
10: gb\_est10:\*  
11: gb\_est11:\*  
12: gb\_est12:\*  
13: gb\_est13:\*  
14: gb\_est14:\*  
15: gb\_est15:\*  
16: gb\_est16:\*  
17: gb\_est17:\*  
18: gb\_est18:\*  
19: gb\_est19:\*  
20: gb\_est20:\*  
21: gb\_est21:\*  
22: gb\_est22:\*  
23: gb\_est23:\*  
24: gb\_est24:\*  
25: gb\_est25:\*  
26: gb\_est26:\*  
27: gb\_est27:\*  
28: gb\_est28:\*  
29: gb\_est29:\*  
30: gb\_est30:\*  
31: gb\_est31:\*  
32: gb\_est32:\*  
33: gb\_est33:\*  
34: gb\_est34:\*  
35: gb\_est35:\*  
36: gb\_est36:\*  
37: gb\_est37:\*  
38: gb\_est38:\*  
39: gb\_est39:\*  
40: gb\_est40:\*  
41: gb\_est41:\*  
42: gb\_est42:\*  
43: gb\_est43:\*

44: em\_esthum10:\*  
45: em\_esthum11:\*  
46: em\_esthum12:\*  
47: em\_esthum13:\*  
48: em\_esthum14:\*  
49: em\_esthum15:\*  
50: em\_esthum16:\*  
51: em\_esthum17:\*  
52: em\_esthum18:\*  
53: em\_esthum19:\*  
54: em\_esthum20:\*  
55: em\_esthum21:\*  
56: em\_esthum22:\*  
57: em\_esthum23:\*  
58: em\_esthum24:\*  
59: em\_esthum25:\*  
60: em\_esthum26:\*  
61: em\_esthum27:\*  
62: em\_esthum28:\*  
63: em\_esthum29:\*  
64: em\_esthum30:\*  
65: em\_esthum31:\*  
66: em\_esthum32:\*  
67: em\_esthum33:\*  
68: em\_esthum34:\*  
69: em\_esthum35:\*  
70: em\_esthum36:\*  
71: em\_esthum37:\*  
72: em\_esthum38:\*  
73: em\_esthum39:\*  
74: em\_esthum40:\*  
75: em\_esthum41:\*  
76: em\_esthum42:\*  
77: em\_esthum43:\*  
78: em\_esthum44:\*  
79: em\_esthum45:\*  
80: em\_esthum46:\*  
81: em\_esthum47:\*  
82: em\_esthum48:\*  
83: em\_esthum49:\*  
84: em\_esthum50:\*  
85: em\_esthum51:\*  
86: em\_esthum52:\*  
87: em\_esthum53:\*  
88: em\_esthum54:\*  
89: em\_esthum55:\*  
90: em\_esthum56:\*  
91: em\_esthum57:\*  
92: em\_esthum58:\*  
93: em\_esthum59:\*  
94: em\_esthum60:\*  
95: em\_esthum61:\*  
96: em\_esthum62:\*  
97: em\_esthum63:\*  
98: em\_esthum64:\*  
99: em\_esthum65:\*  
100: em\_esthum66:\*  
101: em\_esthum67:\*  
102: em\_esthum68:\*  
103: em\_esthum69:\*  
104: em\_esthum70:\*  
105: em\_esthum71:\*  
106: em\_esthum72:\*  
107: em\_esthum73:\*  
108: em\_esthum74:\*  
109: em\_esthum75:\*  
110: em\_esthum76:\*  
111: em\_esthum77:\*  
112: em\_esthum78:\*  
113: em\_esthum79:\*  
114: em\_esthum80:\*  
115: em\_esthum81:\*  
116: em\_esthum82:\*

117: gb\_est148:\*  
118: gb\_est149:\*  
119: gb\_est150:\*  
120: gb\_est151:\*  
121: gb\_est152:\*  
122: gb\_est153:\*  
123: gb\_est154:\*  
124: gb\_est155:\*  
125: gb\_est156:\*  
126: gb\_est157:\*  
127: gb\_est158:\*  
128: gb\_est159:\*  
129: gb\_est160:\*  
130: gb\_est161:\*  
131: gb\_est162:\*  
132: gb\_est163:\*  
133: gb\_est164:\*  
134: gb\_est165:\*  
135: gb\_est166:\*  
136: gb\_est167:\*  
137: gb\_est168:\*  
138: gb\_est169:\*  
139: gb\_est170:\*  
140: gb\_est171:\*  
141: gb\_est172:\*  
142: gb\_est173:\*  
143: gb\_est174:\*  
144: gb\_est175:\*  
145: gb\_est176:\*  
146: gb\_est177:\*  
147: gb\_est178:\*  
148: gb\_est179:\*  
149: gb\_est180:\*  
150: gb\_est181:\*  
151: gb\_est182:\*  
152: gb\_est183:\*  
153: gb\_est184:\*  
154: gb\_est185:\*  
155: gb\_est186:\*  
156: gb\_est187:\*  
157: gb\_est188:\*  
158: gb\_est189:\*  
159: gb\_est190:\*  
160: gb\_est191:\*  
161: gb\_est192:\*  
162: gb\_est193:\*  
163: gb\_est194:\*  
164: gb\_est195:\*  
165: gb\_est196:\*  
166: gb\_est197:\*  
167: gb\_est198:\*  
168: gb\_est199:\*  
169: gb\_est200:\*  
170: gb\_est201:\*  
171: gb\_est202:\*  
172: gb\_est203:\*  
173: gb\_est204:\*  
174: gb\_est205:\*  
175: gb\_est206:\*  
176: gb\_est207:\*  
177: gb\_est208:\*  
178: gb\_est209:\*  
179: gb\_est210:\*  
180: gb\_est211:\*  
181: gb\_est212:\*  
182: gb\_est213:\*  
183: gb\_est214:\*  
184: gb\_est215:\*  
185: gb\_est216:\*  
186: gb\_est217:\*  
187: gb\_est218:\*  
188: gb\_est219:\*  
189: gb\_est220:\*

190: gb\_est110:\*  
191: gb\_est111:\*  
192: gb\_hic:\*  
193: em\_gss\_fun:\*  
194: em\_gss\_hum1:\*  
195: em\_gss\_hum2:\*  
196: em\_gss\_hum3:\*  
197: em\_gss\_hum4:\*  
198: em\_gss\_hum5:\*  
199: em\_gss\_hum6:\*  
200: em\_gss\_hum7:\*  
201: em\_gss\_hum8:\*  
202: em\_gss\_hum9:\*  
203: em\_gss\_inv1:\*  
204: em\_gss\_inv2:\*  
205: em\_gss\_inv3:\*  
206: em\_gss\_other:\*  
207: em\_gss\_pin1:\*  
208: em\_gss\_pin2:\*  
209: em\_gss\_pro:\*  
210: em\_gss\_rod1:\*  
211: em\_gss\_rod2:\*  
212: em\_gss\_rod3:\*  
213: em\_gss\_rod4:\*  
214: em\_gss\_rod5:\*  
215: em\_gss\_vrt1:\*  
216: em\_gss\_vrt2:\*  
217: em\_gss\_vrt3:\*  
218: gb\_gss1:\*  
219: gb\_gss2:\*  
220: gb\_gss3:\*  
221: gb\_gss4:\*  
222: gb\_gss5:\*  
223: gb\_gss6:\*  
224: gb\_gss7:\*  
225: gb\_gss8:\*  
226: gb\_gss9:\*  
227: gb\_gss10:\*  
228: gb\_gss11:\*  
229: gb\_gss12:\*  
230: gb\_gss13:\*  
231: gb\_gss14:\*  
232: gb\_gss15:\*  
233: gb\_gss16:\*  
234: gb\_gss17:\*  
235: gb\_gss18:\*  
236: gb\_gss19:\*  
237: gb\_gss20:\*  
238: gb\_gss21:\*  
239: gb\_gss22:\*  
240: gb\_gss23:\*  
241: gb\_gss24:\*  
242: gb\_gss25:\*  
243: gb\_gss26:\*  
244: gb\_gss27:\*  
245: gb\_gss28:\*  
246: gb\_gss29:\*  
247: gb\_gss30:\*  
248: gb\_gss31:\*  
249: gb\_gss32:\*  
250: gb\_gss33:\*  
251: gb\_gss34:\*  
252: em\_gss\_inv4:\*  
253: em\_gss\_rod6:\*  
254: em\_gss\_rod7:\*  
255: em\_gss\_rod8:\*  
256: gb\_gss35:\*  
257: gb\_gss36:\*  
258: gb\_gss37:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

OS	Homo sapiens.	
XX		Location/Qualifiers
FH	Key	48..689
FT	CDS	/**tag= a
FT		/label= BEER_variant_V101
FT		/product= TGF-beta_binding_protein
XX		
PN	W0200032773-A1.	
XX		
PD	08-JUN-2000.	
XX		
PE	24-NOV-1999;	99WO-US27990.
XX		
PR	27-NOV-1998;	98US-0110283.
XX		
PA	(DARW-) DARWIN DISCOVERY LTD.	
PI	Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paeper BW	
PI	Van Ness J, Winkler DC;	

XX MPI, 2000-412321/35.  
DR P-PSDB; AAY96430.  
XX  
PT Nucleic acids (1) encoding a transforming growth factor beta binding  
PT protein, useful for identifying agents for treating osteopenia,  
PT osteoporosis and fractures  
XX  
PS Claim 1; Page 118-119; 162pp; English.

XX This cDNA encodes a variant human transforming growth factor-beta  
CC (TGF-beta) binding protein designated BEER V101. The encoded protein  
CC comprises a substitution of isoleucine for the wild-type valine at  
CC residue 10. The cDNA and protein may be used for prevention, treatment  
CC and diagnosis of diseases associated with inappropriate BEER expression.  
CC For example, they may be used to treat disorders associated with  
CC decreased TGF-beta BP expression. The cDNA or vectors may be administered  
CC to treat diseases by rectifying mutations or deletions in a patient's  
CC genome that affect the activity of BEER by expressing inactive proteins  
CC or to supplement the patients own production of BEER polypeptides. The  
CC nucleic acids may be used for recombinant production of BEER gene  
CC therapy, antisense therapy, as probes for diagnostic assays and for  
CC functional studies. BEER may be used to raise antibodies and for  
CC identification of BEER modulators. BEER antagonists may be used to  
CC increase bone mineral content for the treatment of disorders such as  
CC osteopenia, osteoporosis, fractures and other disorders associated with  
CC low mineral content.  
XX  
XX Sequence 2301 BP; 569 A; 634 C; 614 G; 484 T; 0 other;

Query Match 100.0%; Score 2301; DB 21; Length 2301;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 agagcgtgtctactggaagtggtggtccctcctgtgctgtacatgacgtccac 60  
DB 1 agagcgtgtctactggaagtggtggtccctcctgtgctgtacatgacgtccac 60  
QY 61 tggcccggtgtctactggtcgtgtgtacacacgcttcgtgaagtggaagcgcaag 120  
DB 61 tggcccggtgtctactggtcgtgtgtacacacgcttcgtgaagtggaagcgcaag 120  
QY 121 ggtgacgagcgttcaagaatgatgacaggaataatcccgagagtcggaagtagcccg 180  
DB 121 ggtgacgagcgttcaagaatgatgacaggaataatcccgagagtagcccg 180  
QY 121 ggtgacgagcgttcaagaatgatgacaggaataatcccgagagtcggaagtagcccg 180  
DB 121 ggtgacgagcgttcaagaatgatgacaggaataatcccgagagtagcccg 180  
QY 181 agctccacacgagcgttgagaacaacaagaacatgacacgagcgaggaagcgagcgcg 240  
DB 181 agctccacacgagcgttgagaacaacaagaacatgacacgagcgaggaagcgagcgcg 240  
QY 241 ctcccccacacaccccttgaagaccaaagacgtgtccagtagacgtcgcgagctgacat 300  
DB 241 ctcccccacacaccccttgaagaccaaagacgtgtccagtagacgtcgcgagctgacat 300  
QY 301 tcacccgctacgtgacacgagtggtggtccgacgacgacgacgacgacgacgacgacg 360  
DB 301 tcacccgctacgtgacacgagtggtggtccgacgacgacgacgacgacgacgacgacg 360  
QY 361 gctccgacgagtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 420  
DB 361 gctccgacgagtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 420  
QY 421 ggcgaactagtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 480  
DB 421 ggcgaactagtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 480  
QY 481 agctggtgtgtccggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 540  
DB 481 agctggtgtgtccggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 540  
QY gcaagtgcaagcgctcaccgcttcacacacagtgtagagtcgaagtagcttcgagacg 600  
DB gcaagtgcaagcgctcaccgcttcacacacagtgtagagtcgaagtagcttcgagacg 600

DB 541 gcaagtgcaagcgctcaccgcttcacacacagtgtagagtcgaagtagcttcgagacg 600  
QY 601 agcgcgtcgtggtcgcgaagaagtggtggtggtggtggtggtggtggtggtggtggtggtg 660  
DB 601 agcgcgtcgtggtcgcgaagaagtggtggtggtggtggtggtggtggtggtggtggtggtg 660  
QY 661 accagcgagcgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 720  
DB 661 accagcgagcgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 720  
QY 721 gcccgagcgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 780  
DB 721 gcccgagcgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 780  
QY 781 attcatgttaaatgctgtgacacacagtggtggtggtggtggtggtggtggtggtg 840  
DB 781 attcatgttaaatgctgtgacacacagtggtggtggtggtggtggtggtggtggtg 840  
QY 841 atcccgagcgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 900  
DB 841 atcccgagcgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 900  
QY 901 gagggaattgagatcacaagacactgagcagcagcccgctctggtggtggtggtggtggtg 960  
DB 901 gagggaattgagatcacaagacactgagcagcagcccgctctggtggtggtggtggtggtg 960  
QY 961 ttgctgtgtccacttcag 1020  
DB 961 ttgctgtgtccacttcag 1020  
QY 1021 agggagcgtgtgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 1080  
DB 1021 agggagcgtgtgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 1080  
QY 1081 cccctgtgacccctgtgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 1140  
DB 1081 cccctgtgacccctgtgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 1140  
QY 1141 caactgtatgagtgtgttctcagtcctgtgctgtgctgacacacacacacacacacacacac 1200  
DB 1141 caactgtatgagtgtgttctcagtcctgtgctgtgctgacacacacacacacacacacacac 1200  
QY 1201 tacacaatctccttcgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 1260  
DB 1201 tacacaatctccttcgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 1260  
QY 1261 taggatctcgagagagactatgtgacatgattccaaagagactccagtggttgaatggg 1320  
DB 1261 taggatctcgagagagactatgtgacatgattccaaagagactccagtggttgaatggg 1320  
QY 1321 cagaggtgag 1380  
DB 1321 cagaggtgag 1380  
QY 1381 caaggtcacttcacgaatctcagatgtgtggtggtggtggtggtggtggtggtggtggtggtg 1440  
DB 1381 caaggtcacttcacgaatctcagatgtgtggtggtggtggtggtggtggtggtggtggtggtg 1440  
QY 1441 caaacagagaaaaaaag 1500  
DB 1441 caaacagagaaaaaaag 1500  
QY 1501 tccctgag 1560  
DB 1501 tccctgag 1560  
QY 1561 oocctccatctcaag 1620  
DB 1561 oocctccatctcaag 1620  
QY 1621 ggtggtgag 1680  
DB 1621 ggtggtgag 1680







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 8, 2001, 23:52:07 ; Search time 438.73 Seconds  
(without alignments)  
913.093 Million cell updates/sec

Title: US-09-668-021-11

Perfect score: 638

Sequence: 1 atgcagccctactactagccccc.....ctggagaagcctactagag 638

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

N.Geneseq\_0601 :  
1: /SIDSI/gcgdata/geneseq/geneseqn/NA1980.DAT : \*  
2: /SIDSI/gcgdata/geneseq/geneseqn/NA1981.DAT : \*  
3: /SIDSI/gcgdata/geneseq/geneseqn/NA1982.DAT : \*  
4: /SIDSI/gcgdata/geneseq/geneseqn/NA1983.DAT : \*  
5: /SIDSI/gcgdata/geneseq/geneseqn/NA1984.DAT : \*  
6: /SIDSI/gcgdata/geneseq/geneseqn/NA1985.DAT : \*  
7: /SIDSI/gcgdata/geneseq/geneseqn/NA1986.DAT : \*  
8: /SIDSI/gcgdata/geneseq/geneseqn/NA1987.DAT : \*  
9: /SIDSI/gcgdata/geneseq/geneseqn/NA1988.DAT : \*  
10: /SIDSI/gcgdata/geneseq/geneseqn/NA1989.DAT : \*  
11: /SIDSI/gcgdata/geneseq/geneseqn/NA1990.DAT : \*  
12: /SIDSI/gcgdata/geneseq/geneseqn/NA1991.DAT : \*  
13: /SIDSI/gcgdata/geneseq/geneseqn/NA1992.DAT : \*  
14: /SIDSI/gcgdata/geneseq/geneseqn/NA1993.DAT : \*  
15: /SIDSI/gcgdata/geneseq/geneseqn/NA1994.DAT : \*  
16: /SIDSI/gcgdata/geneseq/geneseqn/NA1995.DAT : \*  
17: /SIDSI/gcgdata/geneseq/geneseqn/NA1996.DAT : \*  
18: /SIDSI/gcgdata/geneseq/geneseqn/NA1997.DAT : \*  
19: /SIDSI/gcgdata/geneseq/geneseqn/NA1998.DAT : \*  
20: /SIDSI/gcgdata/geneseq/geneseqn/NA1999.DAT : \*  
21: /SIDSI/gcgdata/geneseq/geneseqn/NA2000.DAT : \*  
22: /SIDSI/gcgdata/geneseq/geneseqn/NA2001.DAT : \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	638	100.0	638	21	AAA29058
2	556	87.1	674	21	AAA29059
3	503.2	78.9	642	21	AAA29057
4	502	78.7	2301	21	AAA29055
5	502	78.7	2301	21	AAA29056
6	502	78.7	2329	22	AAA91023
7	500.4	78.4	2301	21	AAA29061
8	500.4	78.4	2301	21	AAA29062
9	500	78.4	642	21	AAA94051
10	427	66.9	35828	21	AAA29063
11	414.2	64.9	532	21	AAA29060

12	367.4	57.6	5680	21	AAA94049	Human DAN/Cerberus
13	367.4	57.6	9301	21	AAA29064	Human TGF-beta bin
14	357.6	56.1	1104	21	AAA94050	Human DAN/Cerberus
15	106	16.6	196	21	AAC05741	Human secreted pro
16	106	16.6	198	20	AAAX0842	Human secreted pro
17	83.2	13.0	954	22	AAC99782	SKN cell cDNA, SE
18	83.2	13.0	962	21	AAZ61645	cDNA encoding mur1
19	83.2	13.0	962	21	AAZ61736	cDNA encoding mur1
20	83.2	13.0	962	22	AAC99578	SKN cell cDNA, SE
21	83.2	13.0	962	22	AAC99669	SKN cell cDNA, SE
22	67.2	10.5	900	18	AAT47661	cDNA encoding smal
23	67.2	10.5	900	19	AAV37735	Human small CGN-11
24	67.2	10.5	1719	21	AAA04498	Human adult retina
25	67.2	10.5	1774	20	AAK07571	Homo sapiens fetal
26	65.6	10.3	1692	21	AAC79467	cDNA sequence of h
27	65.6	10.3	1710	20	AAK84238	cDNA encoding human
28	57	8.9	412	22	AAF66218	Novel human polyu
29	52.4	8.2	71989	21	AAA29349	S. cellulosus DNA
30	47.8	7.5	37856	21	AAA11992	Platenolide syntha
31	47.2	7.4	44377	18	AAT78508	Platenolide syntha
32	47.2	7.4	44377	18	AAT80414	Human adenosine A1
33	46.4	7.3	114955	20	AAK53491	dxax gene of Therm
34	46.2	7.2	2007	19	AAV71089	Thermus thermophil
35	46.2	7.2	2007	20	AAZ30905	Partial human Notc
36	44.4	7.0	5617	19	AAV57163	Human Notch3 cDNA
37	44.4	7.0	8091	19	AAV57001	Human KILMP cDNA
38	44.2	6.9	3930	21	AAZ44744	Monkey MANGO 245 D
39	43.8	6.9	1388	22	AAF28262	Nucleotide sequenc
40	43.2	6.8	58857	21	AAK58471	cDNA encoding oste
41	43	6.7	2075	18	AAV75444	Orf virus genomic
42	43	6.7	4020	18	AAV91361	Human adenosine A1
43	43	6.7	114955	20	AAK53491	S. venezuelae mact
44	42.8	6.7	11220	21	AAZ87298	
45	42.8	6.7	36778	21	AAZ87318	

#### ALIGNMENTS

RESULT 1	
AAA29058	
ID AAA29058 standard; cDNA; 638 BP.	
XX	
AC AAA29058;	
XX	
DT 12-SEP-2000 (first entry)	
XX	
DE Murine TGF-beta binding protein (BEER) cDNA.	
XX	
KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;	
KW BEER; gene therapy; antilense therapy; fracture; bone mineralization; ss.	
XX	
OS Mus musculus.	
XX	
FH key	location/Qualifiers
FT CDS	1..636
FT	/*tag= a
FT	/product= TGF-beta_binding_protein
XX	
PN WO200032773-A1.	
XX	
PD 08-JUN-2000.	
XX	
XX 24-NOV-1999;	99WO-US27990.
XX	
PR 27-NOV-1998;	98US-0110283.
XX	
PA (DARW-) DARWIN DISCOVERY LTD.	
XX	
FI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;	
PI Van Ness J, Winkler DG;	
XX	
DR WPI; 2000-412321/35.	

DR P-PSDB: AAY96432.

XX Nucleic acids (1) encoding a transforming growth factor beta binding  
 PT protein, useful for identifying agents for treating osteopenia,  
 PT osteoporosis and fractures  
 XX  
 PS Claim 1: Page 123; 162pp; English.

XX This cDNA encodes a murine transforming growth factor-beta (TGF-beta)  
 CC binding protein designated mBER. The cDNA and protein may be used for  
 CC prevention, treatment and diagnosis of diseases associated with  
 CC inappropriate BBR expression. For example, they may be used to treat  
 CC disorders associated with decreased TGF-beta BP expression. The cDNA or  
 CC vectors may be administered to treat diseases by rectifying mutations or  
 CC deletions in a patient's genome that affect the activity of BBR by  
 CC expressing inactive proteins or to supplement the patients own production  
 CC of BBR polypeptides. The nucleic acids may be used for recombinant  
 CC production of BBR, gene therapy, antisense therapy, as probes for  
 CC diagnostic assays and for functional studies. BBR may be used to raise  
 CC antibodies and for identification of BBR modulators. BBR antagonists  
 CC may be used to increase bone mineral content for the treatment of  
 CC disorders such as osteopenia, osteoporosis, fractures and other disorders  
 CC associated with low mineral content.

SQ Sequence 638 BP; 114 A; 224 C; 209 G; 91 T; 0 other;

Query Match 100.0%; Score 638; DB 21; Length 638;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-124;  
 Matches 638; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgcagccctcactagcccgctgctcactcactgtgacgcgctctgtgct 60  
 DB 1 atgcagccctcactagcccgctgctcactcactgtgacgcgctctgtgct 60  
 QY 61 gtggagggcgaggggtgtgcaagccttcaggaatgtgccaagaggtatcccaaggcctt 120  
 DB 61 gtggagggcgaggggtgtgcaagccttcaggaatgtgccaagaggtatcccaaggcctt 120  
 QY 121 ggaagatcaccgagccctcctctgagaaacacagacacagacggcgagaaatgga 180  
 DB 121 ggaagatcaccgagccctcctctgagaaacacagacacagacggcgagaaatgga 180  
 QY 181 ggcagacctccaccatcctctatgacgcgaaggtgtgtcgaagtacagctgcgcgag 240  
 DB 181 ggcagacctccaccatcctctatgacgcgaaggtgtgtcgaagtacagctgcgcgag 240  
 QY 241 ctgcactacacccgcttctctgacagaagccatgcgcagcgcgaagccggtacccgag 300  
 DB 241 ctgcactacacccgcttctctgacagaagccatgcgcagcgcgaagccggtacccgag 300  
 QY 301 ttgtgtgtcctccgcaagtgcggcccgcggtgtgtgtcccaacgcatcggcgcggtg 360  
 DB 301 ttgtgtgtcctccgcaagtgcggcccgcggtgtgtgtcccaacgcatcggcgcggtg 360  
 QY 361 aagttgtgtgagccgcaagagcagatttcgctgtcattcccgagatctgctacccgag 420  
 DB 361 aagttgtgtgagccgcaagagcagatttcgctgtcattcccgagatctgctacccgag 420  
 QY 421 cgggtgtcagctgtgtgtcccgaggcgcgcgcgctgcgcgaaggtgcgtcgtgtg 480  
 DB 421 cgggtgtcagctgtgtgtcccgaggcgcgcgcgctgcgcgaaggtgcgtcgtgtg 480  
 QY 481 gctctgtcaagtgtcaagcgtctacccgcttccacaacagctcggaagctcaagacttc 540  
 DB 481 gctctgtcaagtgtcaagcgtctacccgcttccacaacagctcggaagctcaagacttc 540  
 QY 541 gggccgaggaacgcgcgcgcgcgaagaggtcgaagccgcgcgcgcgcgcgcgcgcgcgc 600  
 DB 541 gggccgaggaacgcgcgcgcgcgaagaggtcgaagccgcgcgcgcgcgcgcgcgcgcgc 600  
 QY 601 aaagccaacagcgagcgtggaagacgctactagag 638  
 DB 601 aaagccaacagcgagcgtggaagacgctactagag 638

DB 601 aaagccaacagcgagcgtggaagacgctactagag 638

RESULT 2

AAA29059  
 ID AAA29059 standard; CDNA; 674 BP.

AC AAA29059.

DT 12-SEP-2000 (first entry)

DE Rat TGF-beta binding protein (BBR) CDNA.

KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;  
 KW BBR; gene therapy; antisense therapy; fracture; bone mineralization; ss.

OS Rattus norvegicus.

FH Key Location/Qualifiers

FT CDS 33..674  
 FT /tag= a  
 FT /product= TGF-beta\_binding\_protein

XX MO200032773-A1.

XX 08-JUN-2000.

XX 24-NOV-1999; 99WO-US27990.

XX 27-NOV-1998; 98US-0110283.

XX (DARW-) DARWIN DISCOVERY LTD.

PI Brunow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;

PI Van Ness J, Winkler DG;

DR MPI: 2000-412321/35.

DR P-PSDB: AAY96433.

PT Nucleic acids (1) encoding a transforming growth factor beta binding

PT protein, useful for identifying agents for treating osteopenia,

PT osteoporosis and fractures

XX Claim 1: Page 125; 162pp; English.

XX This cDNA encodes a rat transforming growth factor-beta (TGF-beta)  
 CC binding protein designated mBER. The cDNA and protein may be used for  
 CC prevention, treatment and diagnosis of diseases associated with  
 CC inappropriate BBR expression. For example, they may be used to treat  
 CC disorders associated with decreased TGF-beta BP expression. The cDNA or  
 CC vectors may be administered to treat diseases by rectifying mutations or  
 CC deletions in a patient's genome that affect the activity of BBR by  
 CC expressing inactive proteins or to supplement the patients own production  
 CC of BBR polypeptides. The nucleic acids may be used for recombinant  
 CC production of BBR, gene therapy, antisense therapy, as probes for  
 CC diagnostic assays and for functional studies. BBR may be used to raise  
 CC antibodies and for identification of BBR modulators. BBR antagonists  
 CC may be used to increase bone mineral content for the treatment of  
 CC disorders such as osteopenia, osteoporosis, fractures and other disorders  
 CC associated with low mineral content.

SQ Sequence 674 BP; 128 A; 240 C; 208 G; 98 T; 0 other;

Query Match 87.1%; Score 556; DB 21; Length 674;

Best Local Similarity 92.8%; Pred. No. 1.7e-107;  
 Matches 596; Conservative 0; Mismatches 40; Indels 6; Gaps 1;

QY 1 atgcagccctcactagcccgctgctcactcactgtgacgcgctctgtgct 60  
 DB 33 atgcagctctcactagcccgctgctcactcactgtgacgcgctctgtgct 92  
 QY 61 gtggagggcgaggggtgtgcaagccttcaggaatgtgccaagaggtatcccaaggcctt 120





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 8, 2001, 23:52:21 ; Search time 438.73 Seconds  
(without alignments)  
964.615 Million cell updates/sec

Title: US-09-668-021-13

Perfect score: 674  
Sequence: 1 gagaccgagtgccctcct.....agctgagagaagcctactag 674

Scoring table:  
IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

riched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N.Geneseq\_0601:\*  
1: /SIDSI/gcgdata/geneseq/geneseqn/NA1980.DAT:\*  
2: /SIDSI/gcgdata/geneseq/geneseqn/NA1981.DAT:\*  
3: /SIDSI/gcgdata/geneseq/geneseqn/NA1982.DAT:\*  
4: /SIDSI/gcgdata/geneseq/geneseqn/NA1983.DAT:\*  
5: /SIDSI/gcgdata/geneseq/geneseqn/NA1984.DAT:\*  
6: /SIDSI/gcgdata/geneseq/geneseqn/NA1985.DAT:\*  
7: /SIDSI/gcgdata/geneseq/geneseqn/NA1986.DAT:\*  
8: /SIDSI/gcgdata/geneseq/geneseqn/NA1987.DAT:\*  
9: /SIDSI/gcgdata/geneseq/geneseqn/NA1988.DAT:\*  
10: /SIDSI/gcgdata/geneseq/geneseqn/NA1989.DAT:\*  
11: /SIDSI/gcgdata/geneseq/geneseqn/NA1990.DAT:\*  
12: /SIDSI/gcgdata/geneseq/geneseqn/NA1991.DAT:\*  
13: /SIDSI/gcgdata/geneseq/geneseqn/NA1992.DAT:\*  
14: /SIDSI/gcgdata/geneseq/geneseqn/NA1993.DAT:\*  
15: /SIDSI/gcgdata/geneseq/geneseqn/NA1994.DAT:\*  
16: /SIDSI/gcgdata/geneseq/geneseqn/NA1995.DAT:\*  
17: /SIDSI/gcgdata/geneseq/geneseqn/NA1996.DAT:\*  
18: /SIDSI/gcgdata/geneseq/geneseqn/NA1997.DAT:\*  
19: /SIDSI/gcgdata/geneseq/geneseqn/NA1998.DAT:\*  
20: /SIDSI/gcgdata/geneseq/geneseqn/NA1999.DAT:\*  
21: /SIDSI/gcgdata/geneseq/geneseqn/NA2000.DAT:\*  
22: /SIDSI/gcgdata/geneseq/geneseqn/NA2001.DAT:\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	674	100.0	674	21	AAA29059
2	556	82.5	638	21	AAA29058
3	536.2	79.6	2301	21	AAA29055
4	536.2	79.6	2329	22	AAA91023
5	534.6	79.3	2301	21	AAA29061
6	534.6	79.3	2301	21	AAA29062
7	533	79.1	2301	21	AAA29056
8	530	78.6	642	21	AAA29057
9	525.2	77.9	642	21	AAA94051
10	424.6	63.0	532	21	AAA29060
11	407.4	60.4	35828	21	AAA29063

12	372.6	55.3	5680	21	AAA94049	Human DAN/Cerberus
13	372.6	55.3	9301	21	AAA29064	Human TGF-beta bin
14	364.4	54.1	1104	21	AAA94050	Human DAN/Cerberus
15	119	17.7	196	21	AAC05741	Human secreted pro
16	119	17.7	198	20	AA40842	Human secreted pro
17	82.8	12.3	954	22	AAC99782	Skin cell CDNA, SE
18	82.8	12.3	962	21	AAZ61645	CDNA encoding murI
19	82.8	12.3	962	21	AAZ61736	CDNA encoding murI
20	82.8	12.3	962	22	AAC99578	Skin cell CDNA, SE
21	82.8	12.3	962	22	AAC99669	Skin cell CDNA, SE
22	64	9.5	900	18	AAT47661	CDNA encoding smaI
23	64	9.5	900	19	AAV37735	Human small CCN-1i
24	64	9.5	1719	21	AAA40498	Human adult testis
25	64	9.5	1719	20	AA407571	Human adult testis
26	63.2	9.4	1692	21	AAC79467	CDNA sequence of h
27	63.2	9.4	1710	20	AA484238	CDNA encoding human
28	56.2	8.3	412	22	AA466218	Novel human polynu
29	54.2	8.0	71989	21	AAA29349	Sorangium cellulos
30	54	8.0	37856	21	AAA11992	S. cellulosum DNA
31	49.2	7.3	5617	19	AAV5163	Partial human NOTC
32	49.2	7.3	8091	19	AAV57001	Human Notch3 CDNA.
33	47.2	7.0	1125	22	AAC84259	Signal transductio
34	47	7.0	1380	19	AAV67187	M. luteus salt-res
35	46.6	6.9	674	21	AAA29059	Rat TGF-beta bindi
36	46.6	6.9	1525	21	AAC44071	Zea mays DNA fragm
37	46.2	6.9	2007	19	AAV71085	dnax gene of Therm
38	46.2	6.9	2007	20	AAZ30905	Thermus thermophil
39	46	6.8	11220	21	AAZ87298	S. venezuelae macr
40	46	6.8	36778	21	AAZ87318	S. venezuelae pik
41	46	6.8	37948	21	AAZ87285	S. venezuelae pik
42	46	6.8	38506	21	AAZ5633	Nucleotide sequenc
43	46	6.8	38506	21	AAZ56001	Recombinant cosmid
44	45.4	6.7	18660	21	AAZ58472	Nucleotide sequenc
45	45.4	6.7	44377	18	AAT78508	Platenolide syntha

#### ALIGNMENTS

RESULT 1	
ID AAA29059 standard; CDNA: 674 BP.	
AC AAA29059;	
XX	
DT 12-SEP-2000 (first entry)	
XX	
DE Rat TGF-beta binding protein (BEER) CDNA.	
XX	
KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;	
KM BEER; gene therapy; antisense therapy; fracture; bone mineralization; ss.	
XX	
OS Rattus norvegicus.	
XX	
FH Key	Location/Qualifiers
FT CDS	33..674
FT	/*tag= a
FT	/product= TGF-beta_binding_protein
XX	
PN WO200032773-A1.	
XX	
PD 08-JUN-2000.	
XX	
PF 24-NOV-1999;	99WO-US27990.
XX	
PR 27-NOV-1998;	98US-0110283.
XX	
PA (DARW-) DARWIN DISCOVERY LTD.	
XX	
PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;	
XX	
PI Van Ness J, Winkler DG;	
XX	
DR WPI; 2000-412321/35.	





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 9, 2001, 01:29:34 ; Search time 222.63 Seconds  
(without alignments)  
685.650 Million cell updates/sec

Title: US-09-668-021-13

Perfect score: 674  
Sequence: 1 gagagaccgagtcgcttctct.....agctgagagacgctactag 674

Scoring table: IDENTITY\_NDC  
Gapop 10.0 , Gapext 1.0

Archived: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents.NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCrUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/Backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Query length	ID	Description
1	82.8	12.3	962	US-09-188-930-40	Sequence 40, Appl
2	82.8	12.3	962	US-09-188-930-209	Sequence 209, App
3	64	9.5	900	US-08-468-847B-1	Sequence 1, Appli
4	54.2	8.0	71989	US-09-443-501A-2	Sequence 2, Appli
5	46	6.8	11220	US-09-105-537-32	Sequence 32, Appl
6	46	6.8	36778	US-09-105-537-5	Sequence 5, Appli
7	46	6.8	38506	US-09-320-878-19	Sequence 19, Appl
8	45.4	6.7	44377	US-08-804-227C-7	Sequence 7, Appli
9	45.4	6.7	44377	US-08-804-198-1	Sequence 1, Appli
10	44.2	6.6	2943	US-08-042-747A-7	Sequence 7, Appli
11	42.6	6.3	43280	US-08-804-227C-1	Sequence 1, Appli
12	42.4	6.3	1780	US-08-933-821-5	Sequence 5, Appli
13	42.4	6.3	1780	US-08-960-507-5	Sequence 5, Appli
14	41.6	6.2	33529	US-09-144-085-3	Sequence 3, Appli
15	41	6.1	3380	US-09-162-373-2	Sequence 2, Appli
16	41	6.1	3380	US-09-162-373-2	Sequence 2, Appli
17	40.4	6.0	1154	US-09-467-946-2	Sequence 15, Appl
18	40.2	6.0	1506	US-09-158-767-7	Sequence 8, Appli
19	40.2	6.0	1506	US-09-158-767-8	Sequence 8, Appli
20	40.2	6.0	1506	US-09-158-767-9	Sequence 9, Appli
21	40.2	6.0	2261	US-09-158-767-1	Sequence 1, Appli
22	39.6	5.9	50937	US-09-428-517-1	Sequence 1, Appli
23	39.4	5.8	3624	US-07-951-715A-6	Sequence 6, Appli
24	39.4	5.8	3624	US-08-459-448A-6	Sequence 6, Appli
25	39.4	5.8	3624	US-08-459-595A-6	Sequence 6, Appli
26	39.4	5.8	3624	US-08-459-504B-6	Sequence 6, Appli
27	39.4	5.8	3624	US-08-459-444-6	Sequence 6, Appli

28	39.4	5.8	3624	US-09-053-549-7	Sequence 7, Appli
29	39.4	5.8	8854	US-09-053-549-1	Sequence 1, Appli
30	39.2	5.8	459	US-08-387-942C-9	Sequence 9, Appli
31	39.2	5.8	675	US-08-998-416-1096	Sequence 1096, Ap
32	38.6	5.7	1845	US-07-732-962A-1	Sequence 1, Appli
33	38.6	5.7	1845	PCR-US92-06106-1	Sequence 1, Appli
34	38.6	5.7	2256	US-08-318-826A-5	Sequence 5, Appli
35	38.6	5.7	2256	US-08-318-826A-5	Sequence 5, Appli
36	38.6	5.7	2256	US-08-370-156-1	Sequence 1, Appli
37	38.6	5.7	3016	US-08-814-095-1	Sequence 1, Appli
38	38.6	5.7	3016	US-08-318-826A-7	Sequence 7, Appli
39	38.6	5.7	3016	US-08-370-156-5	Sequence 5, Appli
40	38.6	5.7	3096	US-08-814-095-5	Sequence 5, Appli
41	38.6	5.7	3096	US-08-370-156-3	Sequence 6, Appli
42	38.6	5.7	3096	US-08-814-095-3	Sequence 3, Appli
43	38.6	5.7	35060	US-08-814-095-7	Sequence 7, Appli
44	38.4	5.7	20235	US-07-642-734C-3	Sequence 3, Appli
45	38.4	5.7	20235	US-08-439-009A-3	Sequence 3, Appli

## ALIGNMENTS

RESULT 1  
US-09-188-930-40  
; Sequence 40, Application US/09188930A  
; Patent No. 6150502  
; GENERAL INFORMATION:  
; APPLICANT: Watson, James D.  
; APPLICANT: Strachan, Lorna  
; APPLICANT: Strleman, Matthew  
; APPLICANT: Olmstead, Rene  
; APPLICANT: Murison, James Greg  
; TITLE OF INVENTION: Compositions Isolated From Skin Cells  
; FILE REFERENCE: 11000.1011c1  
; CURRENT APPLICATION NUMBER: US/09/188,930A  
; CURRENT FILING DATE: 1998-11-09  
; NUMBER OF SEQ ID NOS: 348  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 40  
; LENGTH: 962  
; TYPE: DNA  
; ORGANISM: mouse  
US-09-188-930-40

Query Match 12.3%; Score 82.8; DB 3; Length 962;  
Best Local Similarity 53.9%; Pred. No. 1.2e-10;  
Matches 292; Conservative 0; Mismatches 227; Indels 23; Gaps 5;

114 gcttcaagatgagtcgacagaaatccgagcagtcagagagagtcagagtcct 173  
113 gcttcaagatgagtcgacagaaatccgagcagtcagagagagtcagagtcct 172  
174 caggaactagagaaacacacagacatgagcagagcagagcagagaccccccac 233  
173 c--accgagcagacacagcagcagtcagacagcagagagcagagcagtcattc 229  
234 catctc-----tatgacaccaaagcgtgtccgagtcagtcagtcgagagtcactaac 289  
230 agacactgagtcagtcagacacagtcagtcagtcagtcagtcagtcagtcagtc 288  
290 ccgcttgcgttacgagcagcagtcgagcagtcgagcagtcgagcagtcgagtcg 349  
289 caataactattcgagcagcagtcgagcagtcgagcagtcgagcagtcgagtcg 348  
350 gggcagcagtcgagcagcagtcgagcagtcgagcagtcgagcagtcgagtcg 397  
349 ggggagagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcg 408  
398 gaaatggtgagcagcagcagtcgagtcgagtcgagtcgagtcgagtcgagtcg 457

```
Db 409 gttactgagccgagagagctctcaggagtgtgtgtlcaacgacgaagcgcgaccga 468
Qy 458 gggagtgcagctgtgtgtcccccgcgagcgccgcgtctgcgaagtgctgtgt 517
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 469 gaggatccagctgcagtgctcagagacg-c--agcacgcgcacctaacaatcacctgtgt 525
Qy 518 ggcctgtgcagagtgcgaacgcctccgcgtctccaaacacagctgaggtcgaactt 577
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 526 cagcggtgcagagtgcaagaggtacacccgtcagcacacagaggtccagccaaacttga 585
Qy 578 cggacctgcagaccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 637
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 586 aagcgtgtgcgcagcccaagcccgccagcaccacagagagcggaagagagcagcaaatc 645
Qy 638 ca 639
    ||
Db 646 ca 647
```

## RESULT 2

Sequence 209, Application US/09188930A  
Patent No. 6150502

## GENERAL INFORMATION:

APPLICANT: Watson, James D.  
APPLICANT: Strachan, Lorna  
APPLICANT: Sleeman, Matthew  
APPLICANT: Orrust, Rene  
APPLICANT: Morrison, James Greg  
TITLE OF INVENTION: Compositions Isolated From Skin Cells  
FILE REFERENCE: 11000.1011c1  
CURRENT APPLICATION NUMBER: US/09/188,930A  
CURRENT FILING DATE: 1998-11-09  
NUMBER OF SEQ ID NOS: 348  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 209  
LENGTH: 962  
TYPE: DNA  
ORGANISM: Mouse  
US-09-188-930-209

Query Match 12.3%; Score 82.8; DB 3; Length 962;  
Best Local Similarity 53.9%; Pred. No. 1.2e-10;  
Matches 292; Conservative 0; Mismatches 227; Indels 23; Gaps 5;

```
Qy 114 gcttcaagaatgataccacagaataatcattccggagactaagagttaccgaagctctt 173
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 113 gctttaaataatgataccacagaataatccttattcactgtgttaaacctgtlcccgca 172
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 174 caagaactgagagacacacagacatcagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 233
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 173 c---accacagcagcaacagcaccctgaatcagccaggaataatgagcgagcatttgcgt 229
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 234 catcct-----tatgacacaaagacgtgtcagatcagctgtccgcgcgcgcgcgcgcgcgc 289
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 230 agcactgcagctgagtcgaacacagctcagttcaagt-gggctgcagggagacttcggtlccac 288
Qy 290 ccgcttcgtagcagcagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 349
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 289 caaatcatttcgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 348
Qy 350 gggcagtgccgcccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 397
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 349 gggcagtgctgtccctcgtccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 408
Qy 398 gaagtgtgtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 457
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 409 gtactgagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 468
Qy 458 ggcggtagcagctgtgtcccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcg 517
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
Db 469 gaggatccagctgtcagtgctcagcagcgc--agcacgcgcacctaacaataatcacgttgt 525
Qy 518 ggcctgtcgaagtgcagagcgcgtctcaccgcgtctccacaacatctgcagttcaagactt 577
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 526 cagcggtgcgaagtgcagaaggtacacccgtcagcacacagaggttccagcacaacttga 585
Qy 578 cggacctgcagaccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 637
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 586 aagcgtgtgcgcagcccaagcccgccagcaccacagagagcggaagagagccagcaaatc 645
Qy 638 ca 639
    ||
Db 646 ca 647
```

## RESULT 3

US-08-468-847B-1  
Sequence 1, Application US/08468847B  
Patent No. 5780263

## GENERAL INFORMATION:

APPLICANT: Hastings, Gregg A. and Adams, Mark D.  
TITLE OF INVENTION: Human CCN-Like Growth Factor  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,  
ADDRESSEE: CECCHI, STEWART & OLSTEIN  
STREET: 6 BECKER FARM ROAD  
CITY: ROSELAND  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 INCH DISKETTE  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/468,847B  
FILING DATE: 6 June 1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: MULLINS, J.G.  
REGISTRATION NUMBER: 33,073  
REFERENCE/DOCKET NUMBER: 325800-442  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 900 BASE PAIRS  
TYPE: NUCLEIC ACID  
STRANDEDNESS: SINGLE  
TOPOLOGY: LINEAR  
MOLECULE TYPE: CDNA  
US-08-468-847B-1

Query Match 9.5%; Score 64; DB 1; Length 900;  
Best Local Similarity 48.0%; Pred. No. 2.4e-06;  
Matches 258; Conservative 0; Mismatches 265; Indels 15; Gaps 2;

```
Qy 114 gcttcaagaatgataccacagaataatcctccggagctcagagagttaccgaagctctt 173
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 184 GCTTTAAATGATGCCACAGAAATCCTTTATTCACATGTTGTTAACTGTTCCAGCA 243
Qy 174 caggaactagagacacacagacatgaaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 233
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 244 CACCCACACACACACACACGCTTGAATCAAGCCAGAAATGAGGCGACGCACTTTTCAGTAA 303
Qy 234 catccttatgacacaaagacgtgtccgagttacagctgtccgcgcgcgcgcgcgcgcgcgcgcgcgc 293
```

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 8, 2001, 23:52:36 ; Search time 438.73 Seconds  
(without alignments)  
761.388 Million cell updates/sec

Title: US-09-668-021-15

Perfect score: 532  
Sequence: 1 agaatgatccacagagaatc.....accagaagcagccgagccga 532

Scoring table: IDENTITY-NUC  
Gapop 10.0 , Gapext 1.0

Archived: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

N.Geneseq\_0601:\*

- 1: /SIDSI/gcgdata/geneseq/geneseqn/NA1980.DAT:\*
- 2: /SIDSI/gcgdata/geneseq/geneseqn/NA1981.DAT:\*
- 3: /SIDSI/gcgdata/geneseq/geneseqn/NA1982.DAT:\*
- 4: /SIDSI/gcgdata/geneseq/geneseqn/NA1983.DAT:\*
- 5: /SIDSI/gcgdata/geneseq/geneseqn/NA1984.DAT:\*
- 6: /SIDSI/gcgdata/geneseq/geneseqn/NA1985.DAT:\*
- 7: /SIDSI/gcgdata/geneseq/geneseqn/NA1986.DAT:\*
- 8: /SIDSI/gcgdata/geneseq/geneseqn/NA1987.DAT:\*
- 9: /SIDSI/gcgdata/geneseq/geneseqn/NA1988.DAT:\*
- 10: /SIDSI/gcgdata/geneseq/geneseqn/NA1989.DAT:\*
- 11: /SIDSI/gcgdata/geneseq/geneseqn/NA1990.DAT:\*
- 12: /SIDSI/gcgdata/geneseq/geneseqn/NA1991.DAT:\*
- 13: /SIDSI/gcgdata/geneseq/geneseqn/NA1992.DAT:\*
- 14: /SIDSI/gcgdata/geneseq/geneseqn/NA1993.DAT:\*
- 15: /SIDSI/gcgdata/geneseq/geneseqn/NA1994.DAT:\*
- 16: /SIDSI/gcgdata/geneseq/geneseqn/NA1995.DAT:\*
- 17: /SIDSI/gcgdata/geneseq/geneseqn/NA1996.DAT:\*
- 18: /SIDSI/gcgdata/geneseq/geneseqn/NA1997.DAT:\*
- 19: /SIDSI/gcgdata/geneseq/geneseqn/NA1998.DAT:\*
- 20: /SIDSI/gcgdata/geneseq/geneseqn/NA1999.DAT:\*
- 21: /SIDSI/gcgdata/geneseq/geneseqn/NA2000.DAT:\*
- 22: /SIDSI/gcgdata/geneseq/geneseqn/NA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	532	100.0	532	21	AAA29060
2	467.8	87.9	642	21	AAA94051
3	467.8	87.9	2301	21	AAA29055
4	467.8	87.9	2301	21	AAA29061
5	467.8	87.9	2329	22	AAA91023
6	466.2	87.6	2301	21	AAA29056
7	466.2	87.6	2301	21	AAA29052
8	461.4	86.7	642	21	AAA29057
9	424.6	79.8	674	21	AAA29059
10	414.2	77.9	638	21	AAA29058
11	375.2	70.5	5680	21	AAA94049

12	375.2	70.5	9301	21	AAA29064	Human TGF-beta bin
13	364.6	68.5	1104	21	AAA94050	Human DAN/Cerberus
14	332.4	62.5	35828	21	AAA29063	Human TGF-beta bl
15	93	17.5	954	22	AAC99782	Human TGF-beta bl
16	93	17.5	962	21	AA651645	Human TGF-beta bl
17	93	17.5	962	21	AA651736	Human TGF-beta bl
18	93	17.5	962	22	AAC99578	Human TGF-beta bl
19	93	17.5	962	22	AAC99669	Human TGF-beta bl
20	72.2	13.6	900	18	AA147661	Human TGF-beta bl
21	72.2	13.6	900	19	AA147735	Human TGF-beta bl
22	72.2	13.6	1692	21	AAC79467	Human TGF-beta bl
23	72.2	13.6	1710	20	AA840438	Human TGF-beta bl
24	72.2	13.6	1719	21	AAA40498	Human TGF-beta bl
25	72.2	13.6	1774	20	AA807571	Human TGF-beta bl
26	61	11.5	412	22	AA662318	Human TGF-beta bl
27	55.6	10.5	37856	21	AAA11992	Human TGF-beta bl
28	53.2	10.4	114955	20	AA53491	Human TGF-beta bl
29	49.6	9.3	114955	20	AA53491	Human TGF-beta bl
30	49.2	9.2	198	20	AA40842	Human TGF-beta bl
31	48.8	9.2	196	21	AAC05741	Human TGF-beta bl
32	48	9.0	19053	21	AA59146	Human TGF-beta bl
33	48	9.0	19053	21	AA59146	Human TGF-beta bl
34	47.8	9.0	71989	21	AAA29349	Human TGF-beta bl
35	46.6	8.8	1479	21	AA51609	Human TGF-beta bl
36	46.6	8.8	1479	21	AA51609	Human TGF-beta bl
37	45.2	8.5	3861	14	AA031543	Human TGF-beta bl
38	44.8	8.4	2140	15	AA031543	Human TGF-beta bl
39	44.8	8.4	2140	16	AA031543	Human TGF-beta bl
40	44.8	8.4	2140	22	AA624205	Human TGF-beta bl
41	44.4	8.3	1485	21	AA61590	Human TGF-beta bl
42	44.4	8.3	1485	21	AA61590	Human TGF-beta bl
43	44.4	8.3	2277	19	AA13836	Human TGF-beta bl
44	44.4	8.3	2277	19	AA13836	Human TGF-beta bl
45	44.4	8.3	44377	18	AA178508	Human TGF-beta bl

#### ALIGNMENTS

RESULT 1	AAA29060	standard: CDNA: 532 BP.
XX	AC	AAA29060;
XX	DT	12-SEP-2000 (first entry)
DE	XX	Bovine TGF-beta binding protein (BEER) CDNA.
XX	KW	osteopathic: transforming growth factor-beta; TGF-beta: binding protein;
XX	KW	BEER; gene therapy: antisense therapy; fracture; bone mineralization; ss.
XX	OS	Bos taurus.
XX	FT	Key
XX	FT	CDS
XX	FT	Location/Qualifiers
XX	FT	1..530
XX	FT	/*tag= a
XX	FT	/product= TGF-beta_binding-protein
XX	PN	MO200032773-A1.
XX	PD	08-JUN-2000.
XX	XX	
XX	XX	24-NOV-1999;
XX	XX	99WO-US27990.
XX	XX	27-NOV-1998;
XX	XX	98US-0110283.
XX	XX	(DARW-) DARWIN DISCOVERY LTD.
XX	XX	Brunkow ME, Gales DJ, Kovacevich B, Mulligan JT, Paepker BW;
XX	XX	Van Ness J, Winkler DG;

DR WPI: 2000-412321/35.  
P-PSDB; AAT96434.

XX Nucleic acids (1) encoding a transforming growth factor beta binding  
PT protein, useful for identifying agents for treating osteopenia,  
PT osteoporosis and fractures

PS Claim 1: Page 126; 162pp; English.

XX This cDNA encodes a bovine transforming growth factor-beta (TGF-beta)  
CC binding protein designated bBER. The cDNA and protein may be used for  
CC prevention, treatment and diagnosis of diseases associated with  
CC inappropriate BBR expression. For example, they may be used to treat  
CC disorders associated with decreased TGF-beta BP expression. The cDNA or  
CC vectors may be administered to treat diseases by rectifying mutations or  
CC deletions in a patient's genome that affect the activity of BBR by  
CC expressing inactive proteins or to supplement the patient's own production  
CC of BBR polypeptides. The nucleic acids may be used for recombinant  
CC production of BBR, gene therapy, antisense therapy, as probes for  
CC diagnostic assays and for functional studies. BBR may be used to raise  
CC antibodies and for identification of BBR modulators. BBR antagonists  
CC may be used to increase bone mineral content for the treatment of  
CC disorders such as osteopenia, osteoporosis, fractures and other disorders  
CC associated with low mineral content.

SO Sequence 532 BP; 95 A; 199 C; 176 G; 62 T; 0 other;

Query Match 100.0%; Score 532; DB 21; Length 532;  
Best Local Similarity 100.0%; Pred. No. 4.8e-95;  
Matches 532; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 agaatgagcacaagaatcatcccccagcgtgagagctaccacgagcctctgcaagc 60  
DB 1 agaatgagcacaagaatcatcccccagcgtgagagctaccacgagcctctgcaagc 60  
QY 61 tgaacaacaagaacatgaacccgagcgaggaagagagagacccaccaccccttgg 120  
DB 61 tgaacaacaagaacatgaacccgagcgaggaagagagagacccaccaccccttgg 120  
QY 121 agacacaagaagcctcagagctacagctccgagagctcacttaccgcgtacgtacgc 180  
DB 121 agacacaagaagcctcagagctacagctccgagagctcacttaccgcgtacgtacgc 180  
QY 121 agacacaagaagcctcagagctacagctccgagagctcacttaccgcgtacgtacgc 180  
DB 121 agacacaagaagcctcagagctacagctccgagagctcacttaccgcgtacgtacgc 180  
QY 181 atgggagcggtgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 240  
DB 181 atgggagcggtgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 240  
QY 241 cggcgagcgtgctgacaaagcagcagcagcagcagcagcagcagcagcagcagcagc 300  
DB 241 cggcgagcgtgctgacaaagcagcagcagcagcagcagcagcagcagcagcagcagc 300  
QY 301 acttcgctgcatcccgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 360  
DB 301 acttcgctgcatcccgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 360  
QY 361 ggcgagcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 420  
DB 361 ggcgagcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 420  
QY 421 ctgcgttcacaaacagctcagagctcagagcagcagcagcagcagcagcagcagcagc 480  
DB 421 ctgcgttcacaaacagctcagagctcagagcagcagcagcagcagcagcagcagcagc 480  
QY 481 cggagcgagagctgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 532  
DB 481 cggagcgagagctgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 532

RESULT 2  
AAA94051  
ID AAA94051 standard; DNA; 642 BP.  
XX

AC AAA94051;  
XX 15-JAN-2001 (first entry)  
DT Human DAN/Cerberus-related protein 6 (hDCR6) cDNA exons 1 and 4.  
XX  
DE Human DAN/Cerberus-related protein 6 (hDCR6) cDNA exons 1 and 4.  
XX  
KW Human; DNA/Cerberus-related protein 6; hDCR6; morphogenic protein;  
KW antagonist; BMP; cell growth; cell differentiation; bone formation;  
KW gene therapy; ss.  
XX  
XX Homo sapiens.  
OS  
FH Key Location/Qualifiers  
FT CDS 1..642  
FT /tag= a  
FT /product= "hDCR6 #2"  
FT /partial

PN W020005193-A2.  
XX  
XX 21-SEP-2000.  
PD  
PD 02-MAR-2000; 2000WO-US05537.  
XX  
XX 12-MAR-1999; 99US-0124116.  
PR  
XX (REG-) REGENERON PHARM INC.  
XX  
XX Economides AN;  
XX  
XX WPI: 2000-638179/61.  
DR P-PSDB; AAB26106.  
XX  
XX Novel isolated, human DNA/Cerberus related protein 6 which include  
PT natural homologue, and polypeptides comprising DCR6 domain and nucleic  
PT acids encoding the proteins which are useful as probes and primers -  
XX  
XX Claim 2; Fig 3; 40pp; English.

XX The present sequence comprises exons 1 and 4 of the human  
CC DAN/Cerberus-related protein 6 (hDCR6) coding sequence. It was isolated  
CC from a human kidney cDNA library containing exons 1 and 4 of the  
CC sequence. hDCR6 is closely related to the DAN and DCR5 proteins, both of  
CC which act as antagonists of morphogenic proteins such as BMP. It is  
CC possible that the hDCR6 gene and protein can be used as immunogens,  
CC modulators of cell function, growth and differentiation, to reduce  
CC undesirable bone formation, to identify DCR6 binding agents, in  
CC diagnosis, and in gene therapy.

SO Sequence 642 BP; 115 A; 230 C; 211 G; 86 T; 0 other;

Query Match 87.9%; Score 467.8; DB 21; Length 642;  
Best Local Similarity 93.5%; Pred. No. 1.3e-82;  
Matches 500; Conservative 0; Mismatches 32; Indels 3; Gaps 1;

QY 1 agaatgagcacaagaatcatcccccagcgtgagagctaccacgagcctctgcaagc 60  
DB 1 agaatgagcacaagaatcatcccccagcgtgagagctaccacgagcctctgcaagc 60  
QY 61 t--gaacaacaagaacatgaacccgagcgaggaagagagagacccaccacccct 117  
DB 61 t--gaacaacaagaacatgaacccgagcgaggaagagagagacccaccacccct 117  
QY 149 tgaagaacaacaagaacatgaacccgagcgaggaagagagagacccaccacccct 208  
DB 149 tgaagaacaacaagaacatgaacccgagcgaggaagagagagacccaccacccct 208  
QY 178 cgaatggcgagtgccgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 237  
DB 178 cgaatggcgagtgccgagcagcagcagcagcagcagcagcagcagcagcagcagc 237  
QY 269 cgaatggcgagtgccgagcagcagcagcagcagcagcagcagcagcagcagcagc 328  
DB 269 cgaatggcgagtgccgagcagcagcagcagcagcagcagcagcagcagcagcagc 328  
QY 238 gcccgagcgagcgtgctgacaaagcagcagcagcagcagcagcagcagcagcagcagc 297  
DB 238 gcccgagcgagcgtgctgacaaagcagcagcagcagcagcagcagcagcagcagcagc 297





A:Gene: LH-beta





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2001, 15:48:48 ; Search time 31.13 Seconds  
(without alignments)  
234.385 Million cell updates/sec

Title: US-09-668-021-2  
Perfect score: 1167  
Sequence: 1 MOLPLALCLVCLVHTAFRV.....KPRPRARSANQAELENAY 213

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08  
Maximum Match 1008  
Listing first 45 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	94.5	8.1	159	LSHB_MELGA	P45646 meleagris g
2	87	7.5	958	HIG_DROME	Q08101 drosophila
3	85	7.3	180	DAN_HUMAN	P41271 homo sapien
4	83.5	7.2	141	LSHB_TRTIV	O46482 trichosurus
5	83.5	7.2	972	SVL_MYCLE	O50192 mycobacteri
6	82.5	7.1	1480	SLIT_DROME	P24014 drosophila
7	81.5	7.0	870	BCAL_HUMAN	P56945 homo sapien
8	81.5	7.0	969	SVL_MYCTU	P71698 mycobacteri
9	80	6.9	2333	POG_FMDVI	P03306 f genome po
10	79.5	6.8	118	HR23_MOUSE	P42582 mus musculu
11	79	6.8	118	LSHB_PHYCA	P23330 physeter ca
12	79	6.8	178	DAN_RAT	Q06880 rattus norv
13	79	6.8	355	KLF2_HUMAN	O09547 homo sapien
14	78.5	6.7	166	LSHB_COTUA	P45657 coturnix co
15	78	6.7	1	INCE_CHICK	P53352 gallus gall
16	78	6.7	1210	EGFR_HUMAN	P00533 homo sapien
17	77.5	6.6	951	SFRB_HUMAN	Q12872 homo sapien
18	77	6.6	118	LSHB_BALAC	P33086 balaenopter
19	77	6.6	510	ITAI_HUMAN	Q01101 homo sapien
20	77	6.6	20	TPA_HUMAN	P00710 homo sapien
21	76	6.5	398	MUB1_XENLA	P38565 xenopus lae
22	76	6.5	2142	BAT2_HUMAN	P48634 homo sapien
23	75.5	6.5	318	HK25_RAT	O35767 rattus norv
24	75.5	6.5	354	ATK1_HUMAN	O92856 homo sapien
25	75.5	6.5	1436	WC11_BOVIN	P30205 bos taurus
26	75	6.4	141	LSHB_CERST	O77803 ceratotheri
27	75	6.4	992	EBN6_EBV	P03204 Epstein-Bar
28	74	6.3	138	LSHB_MACRU	O46483 macropus ru
29	74	6.3	470	KLF4_HUMAN	O43474 homo sapien
30	74	6.3	474	KLF4_MOUSE	O60793 mus musculu
31	74	6.3	544	H15_DROME	O94890 drosophila
32	74	6.3	2715	TRX2_HUMAN	O94896 homo sapien
33	73.5	6.3	375	CE10_CHICK	P19336 gallus gall

34	73.5	6.3	682	1	SNK_MOUSE	P53351 mus musculu
35	73.5	6.3	682	1	SNK_RAT	O95012 rattus norv
36	73.5	6.3	685	1	SNK_HUMAN	O94973 homo sapien
37	73	6.3	178	1	DAN_MOUSE	O61477 mus musculu
38	73	6.3	191	1	CBX5_HUMAN	P45973 homo sapien
39	73	6.3	301	1	UL49_HSV1	P10233 herpes simp
40	73	6.3	2336	1	POLG_FMDVI	P49303 f genome po
41	72.5	6.2	128	1	LSHB_STRCA	P80664 struthio ca
42	72.5	6.2	507	1	DAF_CAVPO	O60401 cavia porce
43	72.5	6.2	543	1	PKA_STRCO	P54739 streptomyce
44	72	6.2	243	1	TONB_PSEPU	O05613 pseudomonas
45	72	6.2	379	1	CYRE_MOUSE	P18406 mus musculu

## ALIGNMENTS

RESULT 1	LSHB_MELGA	STANDARD;	PRT;	159 AA.
AC	P45646;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	LUTROPIN BETA CHAIN PRECURSOR (LUTEINIZING HORMONE BETA SUBUNIT) (LSH- BETA) (LSH-B) (LH-B).			
GN	LHB.			
OS	Meleagris gallipavo (Common turkey).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Archosauria; Aves; Neognathae; Galliformes; Meleagrididae; Meleagris.			
OX	NCBI_TaxID=9103;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISUE=Pituitary;			
RX	MEDLINE=95290073; PubMed=7772235;			
RA	You S., Foster L.K., Silsby J.L., el Halaoui M.E., Foster D.N.;			
RT	"Sequence analysis of the turkey LH beta subunit and its regulation by gonadotrophin-releasing hormone and prolactin in cultured pituitary cells.";			
RT	J. Mol. Endocrinol. 14:117-129(1995).			
RL	-!- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.			
CC	-!- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN, LUTROPIN, FOLLITROPIN AND GONADOTROPIN.			
CC	-!- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL: L35519; AA74125.1; ALT_INIT.			
DR	HSSP: P01233; IHRP.			
DR	InterPro: IPR000359; -.			
DR	InterPro: IPR001545; -.			
DR	InterPro: IPR002400; -.			
DR	Pfam: PF00007; Cys_knot; 1.			
DR	PRINTS: PR00438; GRCYSKNOR.			
DR	PROSITE: PS00261; GLYCO_HORMONE_BETA_2; 1.			
DR	PROSITE: PS00689; GLYCO_HORMONE_BETA_2; 1.			
KW	Hormone; Signal; Glycoprotein.			
KW	SIGNAL			
FT	CHAIN	1	39	POTENTIAL.
FT	DISULFID	40	159	LUTROPIN BETA CHAIN.
FT	DISULFID	48	96	BY SIMILARITY.
FT	DISULFID	62	111	BY SIMILARITY.
FT	DISULFID	65	149	BY SIMILARITY.
FT	DISULFID	73	127	BY SIMILARITY.
FT	DISULFID	77	129	BY SIMILARITY.

```

QY      62 GGRPHHFFETKVVSYSCRELHFHYVDGRCBARKVPELVYSSGGCGGARLLPNAIGR 121
Db      43 GRRDP-----CRPIVTVAVEKDCCPOCAVTTTACGGICRTNR----EPVIR 85
QY      122 GKWWRPSPDFRCIPDRYRAORVOLL-CPGEAPARRRVRLVASCKCKR 169
Db      86 SPLGRD--POSSCTYGALRYERWALMGCPIDSDPRV-LLPVALSCRCAR 131

RESULT          2
HIG_DROME      STANDARD;          PRT;    958 AA.
009101;
01-NOV-1995 (Rel. 32, Created)
01-NOV-1995 (Rel. 32, Last sequence update)
01-OCT-2000 (Rel. 40, Last annotation update)
LOCOMOTION-RELATED PROTEIN HIKARU GENKI PRECURSOR.
GN         CG
OS         Drosophila melanogaster (Fruit fly).
OC         Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC         Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC         Ephydriidea; Drosophilidae; Drosophila.
OX         NCBI_TaxID=7227;
RN         [1]
RP         SEQUENCE FROM N.A.
RC         TISSUE-Head:
RX         MEDLINE=93213498; PubMed=8461133;
RA         Hoshino M., Matsuzaki F., Nabeshima Y.-I., Hama C.;
RT         "Hikaru genki," a CNS-specific gene identified by abnormal locomotion
RL         in Drosophila, encodes a novel type of protein."
CC         Neuron 10:395-407(1993).
CC         -|- FUNCTION: HAS A ROLE IN THE DEVELOPMENT OF CNS FUNCTIONS INVOLVED
CC         IN LOCOMOTOR ACTIVITY.
CC         -|- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC         -|- ALTERNATIVE PRODUCTS: 4 ISOPFORMS: 1, 2, 3 (SHOWN HERE) AND 4; ARE
CC         PRODUCED BY ALTERNATIVE SPLICING.
CC         -|- TISSUE SPECIFICITY: EXPRESSED IN PCC NEURONS AND NEUROBLASTS
CC         IN THE PROCEPHALIC NEUROGENIC REGION IN THE CENTRAL NERVOUS
CC         SYSTEM.
CC         -|- DEVELOPMENTAL STAGE: MOST ABUNDANT DURING AND/OR AFTER NEURONAL
CC         DIFFERENTIATION AND DURING CELL SPECIFICATION OR AXOGENESIS.
CC         -|- SIMILARITY: CONTAINS 4 SUSHI (SCR) DOMAINS.
CC         -|- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC
CC         THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC         between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC         the European Bioinformatics Institute. There are no restrictions on its
CC         use by non-profit institutions as long as its content is in no way
CC         modified and this statement is not removed. Usage by and for commercial
CC         entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC         or send an email to license@isb-sib.ch).
CC
DB      EMBL; D13884; BAA02984.1; -
DB      EMBL; D13885; BAA02985.1; -
DB      EMBL; D13886; BAA02986.1; -
DB      EMBL; D13887; BAA02987.1; -
DB      HSSP; P10998; IYVC.
DB      Flybase; FBgn0010114; hlg.
DB      InterPro; IPRO00436; -
DB      InterPro; IPRO03006; -
DB      Pfam; PF00067; 19; 1.
DB      Pfam; PF00084; sushi; 5.
DB      GlycoProtein; Alternative splicing; Immunoglobulin domain; Repeat;
DB      Sushi; Signal.

```

FT	SIGNAL	1	31	POTENTIAL.
FT	CHAIN	32	958	LOCOTONIL-RELATED PROTEIN HIKARU GENKI.
FT	DOMAIN	630	709	IG-LIKE C2-TYPE DOMAIN.
FT	DOMAIN	713	769	SUSHI 1.
FT	DOMAIN	772	828	SUSHI 2.
FT	DOMAIN	831	891	SUSHI 3.
FT	DOMAIN	893	952	SUSHI 4.
FT	SIZE	318	320	CELL ATTACHMENT SITE.
FT	DISULFID	714	755	BY SIMILARITY.
FT	DISULFID	741	768	BY SIMILARITY.
FT	DISULFID	773	814	BY SIMILARITY.
FT	DISULFID	800	827	BY SIMILARITY.
FT	DISULFID	832	877	BY SIMILARITY.
FT	DISULFID	863	890	BY SIMILARITY.
FT	DISULFID	894	939	BY SIMILARITY.
FT	DISULFID	922	952	BY SIMILARITY.
FT	CARBOHYD	376	376	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	525	525	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	605	605	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	620	620	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	752	752	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	789	789	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VASAPPLIC	529	553	MISSING (IN ISOFORM 2 AND ISOFORM 4).
FT	VASAPPLIC	892	958	MISSING (IN ISOFORM 1 AND ISOFORM 2).
SO	SEQUENCE	958 AA:	107027 MW:	4161258B5ABC764 CRC64;
Query Match 7.5%; Score 87; DB 1; Length 958;				
Best Local Similarity 19.7%; Pred. No. 2.1;				
Matches 47; Conservative 24; Mismatches 82; Indels 86; Gaps				
Qy	29	FKNDATETIIEELGEYEPPELENN	-----	53
Db	343	FKGDLGLPPESSIGPEPEPLADQNKLQDQYGNSSARVALLMQRKRSRTAGALS	-----	402
Qy	54	-----KTMNRANG-----GRPHHPEFK-DVSEYSCRELHFRVYTDGRCS	-----	96
Db	403	RPKGGSSSKRTTSRKDKGIYDEAGYTFIRHDDPEPEEBEEDVDLQDTE	-----	457
Qy	97	AKPRTLEVCQGQCGP---ARLLPNAIGRCMKMRPSPGDFRCIPDRYARQVOLLCPGGEA	-----	153
Db	458	---VSEIRFGEIGEMODRLCKIRCVKQKWGP-----LCATNEB	-----	495
Qy	154	PRAKRV---LVASCKCKRLTRFHNSLQDGTGAARPOKGRPRPARSAKANOAEI	-----	209
Db	496	DDNGNVAFQPLYSCHVNRIPS-HLLSLYRNISVTPIPPNRGMKTRLSSTLSTNEI	-----	553

RESULT 3

DAN\_HUMAN

ID DAN\_HUMAN STANDARD: PRT: 180 AA.

AC PA4271:

DT 01-FEB-1995 (Rel. 31, Created)

DT 01-FEB-1995 (Rel. 31, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE NEUROBLASTOMA SUPPRESSOR OF TUMORIGENICITY 1 (ZINC FINGER PROTEIN DAN) (N03).

GN NBL1 OR DAN.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI\_TaxID=9606;

RN (1)

RP SEQUENCE FROM N.A.

RC TISSUE=Lung;

RX MEDLINE=94366724; PubMed=8084583;

RA Enomoto H., Ozaki T., Takahashi E., Nomura N., Tabata S.,

RA Takahashi H., Ohnuma N., Tanabe M., Iwai J., Yoshida H., Matsunaga T.,

RA Sakiyama S.,

RT Identification of human DAN gene, mapping to the putative neuroblastoma tumor suppressor locus."

CC Oncogene 9:2788-2791(1994).

CC -1- FUNCTION: POSSIBLE CANDIDATE AS A TUMOR SUPPRESSOR GENE OF



```

RESULT 2
Q9PWB0 PRELIMINARY: PRT: 272 AA.
AC Q9PWB0;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE CERBERUS HOMOLOG.
OS CER.
OC Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinoptera; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=99439862; PubMed=10508582;
RA Zhu L., Marvin M.J., Gardiner A., Laassar A.B., Mercola M., Stern C.D.,
RA Levin M.;
RT "Cerberus regulates left-right asymmetry of the embryonic head and
heart.";
Cur. Biol. 9:931-938(1999).
EMBL: AF139721; AAD51610.1; -
InterPro: IPR000359; -
DR PROSITE: PS01225; CTCK_2; 1.
DR SMART: SM00041; CT: 1; 1224 MW; 36E4C9F719711BCA CRC64;
SQ SEQUENCE 272 AA; 31224 MW; 36E4C9F719711BCA CRC64;

Query Match
Best Local Similarity 23.7%; Score 106; DB 13; Length 272;
Matches 32; Conservative 20; Mismatches 57; Indels 26; Gaps 5;

QY 45 EPPPELENNK---TNRABNGRPRHPETKDYSEYSCRELHFRVYTDGPRSAKPV 100
RC TISSUE=COLON.
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Ota T., Suzuki Y., Odayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RT "NEO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK024848; BAB15026.1; -
SQ SEQUENCE 168 AA; 19320 MW; D5A4E4E81BDFC0E CRC64;

Db 129 EPPYRKDAKKFMDHFMILRKNSASEVYLPRTKNEHQETRLPFSQSAHSECKV-IV 187
QY 101 TELVCSGCGPARLLPNAIGRGKMWPRSGPDR-----CIPDRYRAQRYOLLCPGGEA 153
Db 188 QNNLCFEKCSS-----FHVPGPDRRLYTFCSKCLPRTFSMKHDLNCT-SSV 233
QY 154 PRAKRVLVASCKCK 168
Db 234 PVYAKVMIVECNCE 248

RESULT 3
Q9PUB2 PRELIMINARY: PRT: 272 AA.
AC Q9PUB2;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE CARONTE.
OS CAR.
OC Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinoptera; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Rodriguez Esteban C., Capdevila J., Economides A.N., Pascual J.,
RA Ortiz A., Izpisua Belmonte J.C.;
RT "Caronte, a novel cer-like protein, mediates the establishment of
left-right asymmetry.";
Mol. Cell. 1:1-10(1999).
EMBL: AF179484; AAD55581.1; -
InterPro: IPR000359; -
DR InterPro: IPR001839; -
DR Prodom: PD000357; -; 1.
DR PROSITE: PS01225; CTCK_2; 1.

```

```

DR SMART: SM00041; CT: 1.
SQ SEQUENCE 272 AA; 31201 MW; 0D89729715771BC2 CRC64;

Query Match
Best Local Similarity 23.7%; Score 104; DB 13; Length 272;
Matches 32; Conservative 19; Mismatches 58; Indels 26; Gaps 5;

QY 45 EPPPELENNK---TNRABNGRPRHPETKDYSEYSCRELHFRVYTDGPRSAKPV 100
RC TISSUE=COLON.
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Ota T., Suzuki Y., Odayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RT "NEO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK024848; BAB15026.1; -
SQ SEQUENCE 168 AA; 19320 MW; D5A4E4E81BDFC0E CRC64;

Db 129 EPPYRKDAKKFMDHFMILRKNSASEVYLPRTKNEHQETRLPFSQSAHSECKV-IV 187
QY 101 TELVCSGCGPARLLPNAIGRGKMWPRSGPDR-----CIPDRYRAQRYOLLCPGGEA 153
Db 188 QNNLCFEKCSS-----FHVPGPDRRLYTFCSKCLPRTFSMKHDLNCT-SSV 233
QY 154 PRAKRVLVASCKCK 168
Db 234 PVYAKVMIVECNCE 248

RESULT 4
Q9H772 PRELIMINARY: PRT: 168 AA.
AC Q9H772;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE CDNA: FLJ21195 FIS, CLONE COL00185.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLON.
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Ota T., Suzuki Y., Odayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RT "NEO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK024848; BAB15026.1; -
SQ SEQUENCE 168 AA; 19320 MW; D5A4E4E81BDFC0E CRC64;

Query Match
Best Local Similarity 24.5%; Score 103; DB 4; Length 168;
Matches 45; Conservative 26; Mismatches 65; Indels 48; Gaps 9;

QY 2 QULPLALCVLLVHTAFRVYGGQWQAFKNDAREIIPELGVEPEPPPELENNKTNRAEN 61
Db 4 KLSLSFLVAVLVKVA-----EARKNRP-----GALPSYKDDSSN-----N 41
QY 62 GGRPHHPETKDYSEYSCRELHFRVYTDGPRSAKPVYTELVCSCQ 108
Db 42 SERWQHQLKEVLASSQELVYTERKYLKSDMCKTOPLRQTVSEGCGRS-RTILNFFCYGQ 100
QY 109 CGPARLLPNAIGRGKMWPRSGPDR-----CIPDRYRAQRYOLLCPGGEA-PARKVRLVA 163
Db 101 CN-SFYIPRHV-----KKEEESFQSCAFCKPQRTSVLVELCEPCLDPFFLAKTIQYK 153
QY 164 SKCK 167
Db 154 QCRG 157

RESULT 5
Q9UEM9 PRELIMINARY: PRT: 493 AA.
AC Q9UEM9;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)

```

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2001, 15:46:18 ; Search time 80.28 seconds  
(without alignments)  
160,848 Million cell updates/sec

Title: US-09-668-021-2  
Perfect score: 1167  
Sequence: 1 MQLPLALCVLLVHTAFRV.....KPRPRASAKANQAELENAY 213

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

arched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_0601:\*

- 1: /SIDSR/gcgdata/geneseq/AA1980.DAT:\*
- 2: /SIDSR/gcgdata/geneseq/AA1981.DAT:\*
- 3: /SIDSR/gcgdata/geneseq/AA1982.DAT:\*
- 4: /SIDSR/gcgdata/geneseq/AA1983.DAT:\*
- 5: /SIDSR/gcgdata/geneseq/AA1984.DAT:\*
- 6: /SIDSR/gcgdata/geneseq/AA1985.DAT:\*
- 7: /SIDSR/gcgdata/geneseq/AA1986.DAT:\*
- 8: /SIDSR/gcgdata/geneseq/AA1987.DAT:\*
- 9: /SIDSR/gcgdata/geneseq/AA1988.DAT:\*
- 10: /SIDSR/gcgdata/geneseq/AA1989.DAT:\*
- 11: /SIDSR/gcgdata/geneseq/AA1990.DAT:\*
- 12: /SIDSR/gcgdata/geneseq/AA1991.DAT:\*
- 13: /SIDSR/gcgdata/geneseq/AA1992.DAT:\*
- 14: /SIDSR/gcgdata/geneseq/AA1993.DAT:\*
- 15: /SIDSR/gcgdata/geneseq/AA1994.DAT:\*
- 16: /SIDSR/gcgdata/geneseq/AA1995.DAT:\*
- 17: /SIDSR/gcgdata/geneseq/AA1996.DAT:\*
- 18: /SIDSR/gcgdata/geneseq/AA1997.DAT:\*
- 19: /SIDSR/gcgdata/geneseq/AA1998.DAT:\*
- 20: /SIDSR/gcgdata/geneseq/AA1999.DAT:\*
- 21: /SIDSR/gcgdata/geneseq/AA2000.DAT:\*
- 22: /SIDSR/gcgdata/geneseq/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1167	100.0	213	21	AAB26106 Human DAN/Cerberus
2	1167	100.0	213	21	AAV96429 Human TGF-beta bin
3	1167	100.0	213	22	AAV97589 Human secreted pro
4	1158	99.2	213	21	AAV96436 Human TGF-beta bin
5	1157	99.1	213	21	AAV96430 Human TGF-beta bin
6	1146	98.2	213	21	AAV96431 Verret TGF-beta bi
7	1072	91.9	367	21	AAV96410 Human DAN/Cerberus
8	1050	90.0	213	21	AAV96433 Rat TGF-beta bindi
9	1029	88.2	211	21	AAV96432 Murine TGF-beta bi
10	901.5	77.2	176	21	AAV96434 Bovine TGF-beta bi
11	350.5	30.0	206	21	AAV75981 Murine skin cell p

Result No.	Score	Query Match	Length	DB ID	Description
12	350.5	30.0	206	21	AAV76031 Murine skin cell p
13	350.5	30.0	206	22	AAV55920 Skin cell protein,
14	350.5	30.0	206	22	AAV55970 Skin cell protein,
15	349.5	29.9	206	18	AAV09408 Human small CCN-11
16	349.5	29.9	206	19	AAV58704 Human small CCN-11
17	349.5	29.9	206	20	AAV95711 Homo sapiens fetal
18	349.5	29.9	206	21	AAV10233 Human adult retina
19	266	22.8	50	20	AAV12009 Human 5' EST seque
20	122.5	10.5	116	18	AAV27654 Secreted protein A
21	122.5	10.5	116	18	AAV4090 Human secreted pro
22	118	10.1	23	21	AAV96435 Mutant human TGF-b
23	104.5	9.0	102	20	AAV03225 Amino acid sequenc
24	104.5	9.0	102	21	AAV10277 Human fetal kidney
25	103	8.8	168	21	AAV84014 Amino acid sequenc
26	98	8.4	158	21	AAV51132 Murine cerebral ne
27	96	8.2	184	19	AAV69293 Human b57 protein
28	96	8.2	184	20	AAV42173 Mouse DRM protein
29	96	8.2	184	21	AAV95961 Human gremilin. Ho
30	90	7.7	184	20	AAV42172 Human DRM protein
31	90	7.7	184	20	AAV42174 Rat DRM protein se
32	90	7.7	429	20	AAV42175 EGFP/DRM fusion pr
33	89	7.6	184	19	AAV29727 DAN and b57 protei
34	88	7.5	558	12	AAV13150 T-PA with -ve char
35	87	7.5	391	20	AAV42178 EGFP/DRM fusion pr
36	86.5	7.4	182	19	AAV29726 DAN and b57 protei
37	86.5	7.4	246	21	AAV78333 Herpes simplex vir
38	84.5	7.2	561	12	AAV12342 T-PA with -ve char
39	82.5	7.1	1480	13	AAV25079 Drosophila SLIT pr
40	82	7.0	150	21	AAV43879 Human cancer assoc
41	82	7.0	267	20	AAV96212 Human cerberus lik
42	82	7.0	267	20	AAV86032 Xenopus cerberus pro
43	82	7.0	270	19	AAV41250 Human tPA. Homo s
44	82	7.0	562	6	AAV50342 Tissue plasminogen
45	82	7.0	562	9	AAV82582

#### ALIGNMENTS

RESULT 1

ID AAB26106 standard; Protein: 213 AA.

XX AAB26106;

XX 15-JAN-2001 (first entry)

DE Human DAN/Cerberus-related protein 6 (hDCR6) (exons 1 and 4).

XX Human: DNA/Cerberus-related protein 6; hDCR6; morphogenic protein;

KW antagonist; BMP; cell growth; cell differentiation; bone formation;

KW gene therapy.

XX Homo sapiens.

OS Homo sapiens.

XX WO20005193-A2.

XX 21-SEP-2000.

XX 02-MAR-2000; 2000WO-US05537.

PF 12-MAR-1999; 99US-0124118.

XX (REG-) REGENERON PHARM INC.

XX Economides AN.

PI WPI: 2000-638179/61.

DR N-PSDB: AAA94051.

XX Novel isolated, human DNA/Cerberus related protein 6 which include

PT natural homologue, and polypeptides comprising DCR6 domain and nucleic

PT acids encoding the proteins which are useful as probes and primers

XX Claim 8; Fig 3; 40pp; English.  
PS  
XX The present sequence comprises the amino acid sequence encoded by exons 1  
CC and 4 of the human DAN/Cerberus-related protein 6 (hDCR6) coding  
CC sequence. The coding sequence was isolated from a human kidney cDNA  
CC library containing exons 1 and 4 of the sequence. hDCR6 is closely  
CC related to the DAN and DCR5 proteins, both of which act as antagonists of  
CC morphogenic proteins such as BMP. It is possible that the hDCR6 gene and  
CC protein can be used as immunogens, modulators of cell function, growth  
CC and differentiation, to reduce undesirable bone formation, to identify  
CC DCR6 binding agents, in diagnosis, and in gene therapy.  
XX  
SQ Sequence 213 AA;

Query Match 100.0%; Score 1167; DB 21; Length 213;  
Best Local Similarity 100.0%; Pred. No. 3.4e-104;  
Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQLPLALCLVCLVHTAFRVVGGQWQAFKNDATETIIPELGEPPEPPELNNKTMNRAE 60  
1 mqlplalclvcllvhtafrrvvggwgafkndatellipelgeypppeleennktmnrae 60  
Dy 61 NGRPRPHHPETKDVSEYSCRELHFTRYVTGDCRSAPKPVTELVCSGGCGPARLLPNAIG 120  
61 ngrprphhpctkdvseyscrelhfttryvtgdpcrsakpvtelvcsgcgparllpnaig 120  
QY 121 RGMWMPSCGDFRCIPDRYRAORVOLLCPGGEAPRAKRYRLVASCKCKLITFRHNOSELK 180  
121 rgmwmpscgdfrcipdryraorvllcpggeaprarckryrlvasckcklitrhngselk 180  
Dy 121 rgmwmpscgdfrcipdryraorvllcpggeaprarckryrlvasckcklitrhngselk 180  
QY 181 DFGTEAARPOKGRKPRPARSAKANOAELENAY 213  
181 dfgteaarpqkgrkprparsakangoelenay 213  
Dy 181 dfgteaarpqkgrkprparsakangoelenay 213

RESULT 2  
AAV96429  
ID AAV96429 standard; Protein; 213 AA.  
XX  
AC AAV96429;  
XX  
DT 12-SEP-2000 (first entry)  
XX  
DE Human TGF-beta binding protein (BEER).  
XX  
KM osteopathic; transforming growth factor-beta; TGF-beta; binding protein;  
KM BEER; gene therapy; antisense therapy; fracture; bone mineralization.  
XX

XX Homo sapiens.

PN WO200032773-A1.

PD 08-JUN-2000.

PE 24-NOV-1999; 99WO-US27990.

PR 27-NOV-1998; 98US-0110283.

PA (DARW-) DARWIN DISCOVERY LTD.

PI Brunhrow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;  
PI Van Ness J, Winkler DG;

XX WPI; 2000-412321/35.

DR N-PSDB; AAA29055.

XX Nucleic acids (I) encoding a transforming growth factor beta binding  
PT protein, useful for identifying agents for treating osteopenia,  
PT osteoporosis and fractures

XX Claim 2; Page 116; 162pp; English.

XX This shows the human transforming growth factor-beta (TGF-beta)  
CC binding protein designated hBEER. The cDNA and protein may be used for  
CC prevention, treatment and diagnosis of diseases associated with  
CC inappropriate BEER expression. For example, they may be used to treat  
CC disorders associated with decreased TGF-beta BP expression. The cDNA or  
CC vectors may be administered to treat diseases by rectifying mutations or  
CC deletions in a patient's genome that affect the activity of BEER by  
CC expressing inactive proteins or to supplement the patient's own production  
CC of BEER polypeptides. The nucleic acids may be used for recombinant  
CC production of BEER, gene therapy, antisense therapy, as probes for  
CC diagnostic assays and for functional studies. BEER may be used to raise  
CC antibodies and for identification of BEER modulators. BEER antagonists  
CC may be used to increase bone mineral content for the treatment of  
CC disorders such as osteopenia, osteoporosis, fractures and other  
CC disorders associated with low mineral content.  
XX  
SQ Sequence 213 AA;

Query Match 100.0%; Score 1167; DB 21; Length 213;  
Best Local Similarity 100.0%; Pred. No. 3.4e-104;  
Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQLPLALCLVCLVHTAFRVVGGQWQAFKNDATETIIPELGEPPEPPELNNKTMNRAE 60  
1 mqlplalclvcllvhtafrrvvggwgafkndatellipelgeypppeleennktmnrae 60  
Dy 61 NGRPRPHHPETKDVSEYSCRELHFTRYVTGDCRSAPKPVTELVCSGGCGPARLLPNAIG 120  
61 ngrprphhpctkdvseyscrelhfttryvtgdpcrsakpvtelvcsgcgparllpnaig 120  
Dy 61 ngrprphhpctkdvseyscrelhfttryvtgdpcrsakpvtelvcsgcgparllpnaig 120  
QY 121 RGMWMPSCGDFRCIPDRYRAORVOLLCPGGEAPRAKRYRLVASCKCKLITFRHNOSELK 180  
121 rgmwmpscgdfrcipdryraorvllcpggeaprarckryrlvasckcklitrhngselk 180  
Dy 121 rgmwmpscgdfrcipdryraorvllcpggeaprarckryrlvasckcklitrhngselk 180  
QY 181 DFGTEAARPOKGRKPRPARSAKANOAELENAY 213  
181 dfgteaarpqkgrkprparsakangoelenay 213  
Dy 181 dfgteaarpqkgrkprparsakangoelenay 213

RESULT 3  
AAV97589  
ID AAV97589 standard; Protein; 213 AA.  
XX  
AC AAV97589;  
XX  
DT 05-APR-2001 (first entry)  
XX  
DE Human secreted protein PRO7476.  
XX  
KM Secreted protein; human; PRO protein; neoplastic cell growth; tumour;  
KM proliferation; leukaemia; lymphoid malignancy; inflammatory disorder;  
KM angiogenic disorder; immunologic disorder; PRO7476.  
XX

XX Homo sapiens.

PN WO200075317-A2.

PD 14-DEC-2000.

PE 15-MAY-2000; 2000MO-US13358.

PR 09-JUN-1999; 99US-0138385.

PR 20-JUL-1999; 99US-0144790.

PR 03-AUG-1999; 99US-0146843.

PR 10-AUG-1999; 99US-0148188.

PR 17-AUG-1999; 99US-0149320.

PR 17-AUG-1999; 99US-0149327.

PR 17-AUG-1999; 99US-0149396.

PR 20-AUG-1999; 99US-0150114.

PR 31-AUG-1999; 99US-0151700.

PR 31-AUG-1999; 99US-0151734.











GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2001, 16:01:12 ; Search time 31.13 Seconds  
(without alignments)  
234.385 Million cell updates/sec

Title: US-09-668-021-6

Perfect score: 1165  
Sequence: 1 MQLPLALCLICLLVHTAFRV.....KPPRRASAKANQAELEENAY 213

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

arched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	94.5	8.1	159	LSHB_MELGA	P45646 meleagris g
2	85	7.3	180	DAN_HUMAN	P41271 homo sapien
3	84.5	7.3	972	SVL_MYCLE	Q50192 mycobacteri
4	83.5	7.2	141	LSHB_TRIIV	O46482 trichosurus
5	81.5	7.0	1480	SLIT_DROME	P24014 drosophila
6	80.5	6.9	969	SVL_MYCTU	P71698 mycobacteri
7	80	6.9	510	SLIT_HUMAN	O01101 homo sapien
8	79.5	6.8	318	HK25_MOUSE	P42582 mus musculu
9	79	6.8	118	LSHB_PRYCA	P25330 physeter ca
10	79	6.8	178	DAN_RAT	O06880 rattus norv
11	79	6.8	243	TOMB_PSEPU	O05613 pseudomonas
12	79	6.8	355	KLP2_HUMAN	Q95W33 homo sapien
13	79	6.8	870	BCAL_HUMAN	P56945 homo sapien
14	79	6.8	2333	POLG_FMDVI	P03306 f genome po
15	78.5	6.7	166	LSHB_COTVA	P45657 coturnix co
16	78	6.7	958	HIG_DROME	O09101 drosophila
17	78	6.7	1210	EGFR_HUMAN	P00533 homo sapien
18	77.5	6.7	951	SFRB_HUMAN	Q12872 homo sapien
19	77.5	6.7	2142	BAT2_HUMAN	P48634 homo sapien
20	77	6.6	118	LSHB_BALAC	P33088 balaenopter
21	77	6.6	877	INCE_CHICK	P53352 gallus gall
22	77	6.6	992	EBN6_EBV	P03304 Epstein-Bar
23	76.5	6.6	354	ATH1_HUMAN	Q92558 homo sapien
24	76	6.5	112	PSC2_RAT	P02781 rattus norv
25	76	6.5	398	MUB1_XENLA	P38565 xenopus lae
26	76	6.5	562	TPA_HUMAN	P00750 homo sapien
27	75.5	6.5	318	HK25_RAT	O35767 rattus norv
28	75.5	6.5	543	PKA1_STRCO	P54739 streptomyce
29	75.5	6.5	1436	MC11_BOVIN	P30305 bos taurus
30	75	6.4	141	LSHB_CERST	O77835 ceratotheri
31	74	6.4	138	LSHB_MACRU	O46483 macropus ru
32	74	6.4	470	KLP4_HUMAN	O43474 homo sapien
33	74	6.4	544	H15_DROME	O94890 drosophila

34	74	6.4	2715	1	TRX2_HUMAN	Q9um66 homo sapien
35	73.5	6.3	301	1	UL49_HSV11	P10233 herpes simp
36	73.5	6.3	682	1	SNK_MOUSE	P53331 mus musculu
37	73.5	6.3	682	1	SNK_MOUSE	O9-012 rattus norv
38	73.5	6.3	685	1	SNK_HUMAN	O9nyy3 homo sapien
39	73	6.3	178	1	DAN_MOUSE	O61477 mus musculu
40	73	6.3	191	1	CBX5_HUMAN	P45973 homo sapien
41	73	6.3	972	1	POLS_IPNVN	P22495 infectious
42	72.5	6.2	128	1	LSHB_STRCA	P80664 struthio ca
43	72.5	6.2	375	1	CE10_CHICK	P19336 gallus gall
44	72.5	6.2	755	1	RRE1_HUMAN	Q92766 homo sapien
45	72	6.2	379	1	CYR6_MOUSE	P18406 mus musculu

## ALIGNMENTS

RESULT 1	LSHB_MELGA	STANDARD;	PRT;	159 AA.
ID	LSHB_MELGA			
AC	P45646;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	LUTROPIN BETA CHAIN PRECURSOR (LUTRINIZING HORMONE BETA SUBUNIT) (LSH- BETA) (LSH-B) (LH-B).			
GN	LHB.			
OS	Meleagris gallipavo (Common turkey).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Archosauria; Aves; Neognathae; Galliformes; Meleagrididae; Meleagris.			
OX	NCBI_TaxID=9103;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Plutary;			
RX	MEDLINE=95290073; PubMed=7772235;			
RA	You S., Foster L.K., Silsby J.L., el Halawani M.E., Foster D.N.;			
RT	"Sequence analysis of the turkey LH beta subunit and its regulation by gonadotrophin-releasing hormone and prolactin in cultured pituitary cells.";			
RT	J. Mol. Endocrinol. 14:117-129(1995).			
RL	-1- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.			
CC	-1- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN, LUTROPIN, FOLLITROPIN AND GONADOTROPIN.			
CC	-1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).			
CC	-----			
CC	EMBL; L35519; AAA74125.1; ALT_INIT.			
DR	HSSP; P01233; IHRP.			
DR	InterPro; IPR000359; -.			
DR	InterPro; IPR001545; -.			
DR	InterPro; IPR002400; -.			
DR	Pfam; PF00007; Cys_knot; 1.			
DR	PRINTS; PR00438; GFCYSKNOT.			
DR	PROSITE; PS00261; GLYCO_HORMONE_BETA_1; 1.			
DR	PROSITE; PS00689; GLYCO_HORMONE_BETA_2; 1.			
FW	Hormone; Signal; Glycoprotein.			
FT	SIGNAL 1 39			
FT	CHAIN 40 159			
FT	DISULFID 48 96			
FT	DISULFID 62 111			
FT	DISULFID 65 149			
FT	DISULFID 73 127			
FT	DISULFID 77 129			
FT	DISULFID 77 129			

FT DISULFID 132 139 BY SIMILARITY.  
 FT CARBOHYD 52 52 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 159 AA; 16285 MW; 52850C8C879653C6 CRC64;

Query Match 8.1%; Score 94.5; DB 1; Length 159;  
 Best Local Similarity 29.4%; Pred. No. 0.051;  
 Matches 32; Conservative 7; Mismatches 49; Indels 21; Gaps 5;

QY 62 GGRPHHPTKDVSEYSCRELHFTYVTDGPCRSAPYTELVCSGCGPARLLPNAIGR 121  
 DB 43 GGRPP-----CRPINVTVAVEKDEPOCMATTTACGGYCTR-----EPYR 85  
 QY 122 GKRRRSGPFCIPDPYRRAQRYOLL-CPEGAPRARKRYLVASCCKR 169  
 DB 86 SPLGRP--PQSSCTGYALRYRWALWCGPISDPKV-LIPVALSCRCAR 131

RESULT 2  
 DAN\_HUMAN  
 ID DAN\_HUMAN STANDARD; PRT; 180 AA.  
 P1271;

01-FEB-1995 (Rel. 31, Created)  
 01-FEB-1995 (Rel. 31, Last sequence update)  
 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE NEUROBLASTOMA SUPPRESSOR OF TUMORIGENICITY 1 (ZINC FINGER PROTEIN DAN) (N03)  
 GN NBL1 OR DAN.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC TISSUE=Lung;  
 RX MEDLINE=94366724; PubMed=8084583;  
 RA Enomoto H., Ozaki T., Takahashi E., Nomura N., Tabata S.,  
 RA Takahashi H., Ohnuma N., Tanabe M., Iwai J., Yoshida H., Matsunaga T.,  
 RA Sakiyama S.;  
 RT "Identification of human DAN gene, mapping to the putative  
 neuroblastoma tumor suppressor locus.";

RL Oncogene 9:2785-2791(1994).  
 CC -1- FUNCTION: POSSIBLE CANDIDATE AS A TUMOR SUPPRESSOR GENE OF  
 NEUROBLASTOMA. MAY PLAY AN IMPORTANT ROLE IN PREVENTING CELLS  
 FROM ENTERING THE FINAL STAGE (G1/S) OF THE TRANSFORMATION  
 PROCESS.  
 CC -1- TISSUE SPECIFICITY: MOST ABUNDANT IN NORMAL LUNG AND MENINGIOMA.  
 CC -1- DISEASE: THE LOSS OF DAN GENE BY MUTATION IS POSSIBLY THE CAUSE OF  
 THE DEVELOPMENT AND/OR PROGRESSION OF HUMAN NEUROBLASTOMA.  
 CC -1- SIMILARITY: HIGH, TO OTHER MAMMALIAN DAN.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC  
 CC EMBL; D28124; BAA05671.1; .  
 DR MIM; 600613; .  
 KM Anti-oncogene; DNA-binding; Zinc-finger.  
 FT 2N FING 81 101 CA-TYPE (POTENTIAL).  
 FT DOMAIN 141 171 PRO-RICH  
 SQ SEQUENCE 180 AA; 19277 MW; 15233229DB655865 CRC64;

Query Match 7.3%; Score 85; DB 1; Length 180;  
 Best Local Similarity 26.0%; Pred. No. 0.45;  
 Matches 33; Conservative 15; Mismatches 61; Indels 18; Gaps 7;

QY 80 CRELHFTYVTDGPCRSAPYTELVCSGCGPARLLPNAIGRKKWRPSGPDF-----RC 134  
 . . . . .

DB 34 CEAKNTQIVGHSGC-EAKSIQNRACLGQCF-SYSPVNTF-----POSTESLVHCDSC 84  
 QY 135 IPDRYRARRVOLLCGPG-GEAPRARK-VRLVASCCKRLTRFNHNSLKPF--GTEAARPO 190  
 DB 85 MPASQMEIYTLCEPGHGEVPRVDKLVKTLHSCQACQCKEPSEHLSYVGGEDGPGSQ 144  
 QY 191 GKRRPR 197  
 DB 145 GGRHNP 151

RESULT 3  
 SYL\_MYCLE  
 ID SYL\_MYCLE STANDARD; PRT; 972 AA.  
 AC Q50192;

DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE LEUCYL-TRNA SYNTHETASE (EC 6.1.1.4) (LEUCINE--TRNA LIGASE) (LEURS).  
 GN LEUS.  
 OS Mycobacterium leprae.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacteriaceae; Mycobacterium.  
 OX NCBI\_Taxid=1769;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RA Eigmeier K., Garner T., de Rossi E., Feil H., Cole S.T.;  
 RL Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: ATP + L-LEUCINE + TRNA(LEU) = AMP +  
 PYROPHOSPHATE + L-LEUCYL-TRNA(LEU).  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC  
 CC EMBL; Y14967; CAA75192.1; .  
 DR EMBL; Z70722; CAA94727.1; .  
 DR Interpro: IPR001412; .  
 DR Interpro: IPR002300; .  
 DR Interpro: IPR002302; .  
 DR Pfam: PF00133; tRNA-synL.1;  
 DR PRINTS: PR00985; tRNAsynL1EU.  
 DR PROSITE: PS00178; AA-TRNA\_LIGASE\_1; 1  
 KM Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.  
 FT SITE 80 89 "HIGH" REGION.  
 FT SITE 741 745 "KMSKS" REGION.  
 FT BINDING 744 744 ATP (BY SIMILARITY).  
 FT CONFLICT 659 668 GGEHAYLH. -> AEPNRCYIC (IN CAA94727).  
 SQ SEQUENCE 972 AA; 108407 MW; 65286839DA51AFAC CRC64;

Query Match 7.3%; Score 84.5; DB 1; Length 972;  
 Best Local Similarity 26.0%; Pred. No. 3;  
 Matches 38; Conservative 15; Mismatches 76; Indels 17; Gaps 6;

QY 16 TAFVVGSGOMAFKNDATETIRELGEYRPPRELNNKTNMAENG-----GRPHNPE 71  
 DB 106 TGHVNLAMGDFAGLPAEDYAMQGTGHHRIITEA--NYYNRRHQGLRGLGHDSSRRFS 163  
 QY 72 TKDVSEYSCRELHFTYVTDGPCRSAPYTELVCSGCGPARLLPNAIGRKKWRPS 128  
 DB 164 TTDFEYFKMTQWIFLQYYNMFVAAWKARPIAELIAFDSSGRRLLVD--GR-DWATLS 219  
 QY 129 GPDFRCIPDR-----YRQRYOLLCPG 150  
 DB 220 AGERADVIDNCRLVYRADSVVNMCPG 245

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2001, 16:00:34 ; Search time 89.34 Seconds

(without alignments)  
315.435 Million cell updates/sec

Title: US-09-668-021-6

Perfect score: 1165

Sequence: 1 MOLPLALICLLVHTAFRV.....KPRPRASAKANQALENNY 213

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Archived: 425026 segs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

SPTREMBL\_16:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organella:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_unclassified:\*  
13: sp\_vertebrate:\*  
14: sp\_virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	270	23.2	134	4 Q9Y3U3	Q9Y3U3 homo sapien
2	108	9.3	168	4 Q9H772	Q9H772 homo sapien
3	106	9.1	272	13 Q9PWB0	Q9PWB0 gallus gall
4	104	8.9	272	13 Q9PWB0	Q9PWB0 gallus gall
5	103	8.8	168	11 Q9H873	Q9H873 mus musculu
6	98.5	8.5	493	4 Q9UEM9	Q9UEM9 homo sapien
7	94.5	8.1	493	4 Q9H876	Q9H876 homo sapien
8	93.5	8.0	493	4 Q9H876	Q9H876 homo sapien
9	93	8.0	1114	11 Q9JLW7	Q9JLW7 mus musculu
10	92	7.9	184	4 Q60565	Q60565 homo sapien
11	91	7.8	215	8 Q9MFB6	Q9MFB6 beta vulgar
12	90	7.7	184	11 Q9JLW7	Q9JLW7 mus musculu
13	90	7.6	492	11 Q9JLW7	Q9JLW7 mus musculu
14	88.5	7.5	184	13 Q73755	Q73755 gallus gall
15	87.5	7.4	182	13 Q73754	Q73754 xenopus lae
16	86.5	7.1	661	5 Q9V708	Q9V708 drosophila
17	82.5	7.1	791	11 Q70368	Q70368 talus norv
18	82	7.0	270	13 P70041	P70041 xenopus lae

20	82	7.0	988	6 Q97867	Q97867 sus scrofa
21	81.5	7.0	1480	5 Q9V7E8	Q9V7E8 drosophila
22	81.5	7.0	1504	5 Q9XYV4	Q9XYV4 drosophila
23	81.5	7.0	1504	5 Q9V7E9	Q9V7E9 drosophila
24	81	7.0	355	4 Q9UKR6	Q9UKR6 homo sapien
25	80.5	6.9	267	4 Q9S813	Q9S813 homo sapien
26	80	6.9	712	4 Q9Y557	Q9Y557 homo sapien
27	80	6.9	720	4 Q9N1Z3	Q9N1Z3 homo sapien
28	80	6.9	866	5 P91426	P91426 caenorhabdi
29	80	6.9	1440	5 Q20204	Q20204 caenorhabdi
30	80	6.9	2531	5 Q21980	Q21980 caenorhabdi
31	79.5	6.8	500	10 Q9M6C2	Q9M6C2 trifolium r
32	79.5	6.8	1081	4 Q76065	Q76065 homo sapien
33	79.5	6.8	1122	2 Q85018	Q85018 mycobacteri
34	79.5	6.8	1181	5 Q9V795	Q9V795 drosophila
35	79	6.8	220	10 Q9P995	Q9P995 oryza sativ
36	79	6.8	329	2 Q9W360	Q9W360 acetobacter
37	79	6.8	355	4 Q9U0S5	Q9U0S5 homo sapien
38	79	6.8	542	2 Q68872	Q68872 myxococcus
39	78.5	6.7	240	10 Q9SNN5	Q9SNN5 oryza sativ
40	78.5	6.7	365	2 Q52730	Q52730 rhizobium e
41	78.5	6.7	482	14 Q71654	Q71654 human immun
42	78.5	6.7	1096	10 Q9XE24	Q9XE24 oryza sativ
43	78.5	6.7	1138	10 Q9ZPF9	Q9ZPF9 arabidopsis
44	78	6.7	224	14 Q9DUD2	Q9DUD2 foot-and-mo
45	78	6.7	224	14 Q9DUD1	Q9DUD1 foot-and-mo

## ALIGNMENTS

RESULT 1  
ID Q9Y3U3 PRELIMINARY: PRT: 134 AA.  
AC -Q9Y3U3;  
DT 01-NOV-1999 (TREMBLrel. 12, Created)  
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
DE HYPOTHETICAL 15.3 KDA PROTEIN (FRAGMENT).  
GN DKFZP564D206.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=BRAIN;  
RA Wambutt R., Heuber D., Mewes H.W., Gassenhuber J., Wiemann S.;  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AL050024; CAB43243.1; .  
DR InterPro: IPR000359; .  
DR PROSITE: PS01225; CTCK\_2; 1.  
KW Hypothetical protein.  
FT NON\_TER  
SQ SEQUENCE 134 AA; 15324 MW; A0B7A8650D2EE6F1 CRC64;

Query Match 23.2%; Score 270; DB 4; Length 134;  
Best local similarity 44.2%; Pred. No. 1.7e-19;  
Matches 57; Conservative 23; Mismatches 43; Indels 6; Gaps 4;

QY 80 CRELHTRVVDGPGCRSAKPYTELVCSGGCGPARLDPNAIGRG---KMW-RPSGDPFCTI 135  
DB 3 CRLRSTKTYISDGOCTSIISPLKEIVCAECILLPVLPMWIGGCGTKYKRSSEBWRV 62  
QY 136 PDYRQVROVLCPGGEPRARAKVLAASCKRLTFRNQSELEKDFGEARPOKGRK 195  
DB 63 NDKTRFRQLQDQGDG-STRTYKITYVTACKCKRYTPHNESSHNPESSPAKPYQHIRE 121  
QY 196 RPRR-RSAK 203  
DB 122 RRRASKSK 130

```

RESULT 2
Q9H772 PRELIMINARY; PRT: 168 AA.
ID Q9H772;
AC Q9H772;
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
DE CUNA: FLJ21195 FIS, CLONE COL00185.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLON;
RA Kawabata A., Hiki T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
Okitani R., Ota T., Suzuki Y., Odayashi M., Nishi T., Shibahara T.,
Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RT "NEDO human cDNA sequencing project."
RL Submitted (AUG-2000) to the EMBL/Genbank/DBJ databases.
EMBL: AF024848; BAB15026.1;
SEQUENCE 168 AA; 19320 MW; D5A4E4EB18BF8C0E CRC64;

Query Match 9.3%; Score 108; DB 4; Length 168;
Best Local Similarity 24.5%; Pred. No. 0.0032;
Matches 45; Conservative 27; Mismatches 64; Indels 48; Gaps 9;

QY 2 QLPALCLICLVHTAFRVVGGQGAFAKNDATETIRELGEYPPPELENNKTMNRAEN 61
Db 4 KLSLTLVAVLVKVA-----EARKN-----RPAGAIPSPYKDGSSN-----N 41

QY 62 GGRPRHPFETKDYSEYSESCHELFHTRYVYDGCGRSAKRY 108
Db 42 SERKQKQKEVLAASQEAVALVTERKYLSDMKCTDPLQVYSEGCXS-RTILNRFCTYQ 100

QY 109 CGPARLLPNAIGRGKWMRPSGPDFR-----CIPDRYRAQRYOLLCPGGEAP-RAKRYVLA 163
Db 101 CN-SFYIPRHV-----KKEESPQSCAFCKPQRYVTSVLVELCPGIDPFFRLKTIQKVA 153

QY 164 SCKK 167
Db 154 QCRC 157

RESULT 3
Q9PWB0 PRELIMINARY; PRT: 272 AA.
ID Q9PWB0;
AC Q9PWB0;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DE CERBERUS HOMOLOG.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=99439862; PubMed=10508582;
RA Zhu L., Marvin M.J., Gardiner A., Lassar A.B., Mercola M., Stern C.D.,
Levin M.;
RT "Cerberus regulates left-right asymmetry of the embryonic head and
heart."
RL Curr. Biol. 9:931-938(1999).
DR EMBL: AF139721; AAD51610.1;
DR InterPro: IPR000359;
DR PROSITE: PS01225; CTCK_2; 1.
DR SMART: SM00041; CT: 1; 1224 MW; 36E4C9F19711BCA CRC64;
SQ SEQUENCE 272 AA; 31224 MW; 36E4C9F19711BCA CRC64;

```

```

Query Match 9.1%; Score 106; DB 13; Length 272;
Best Local Similarity 23.7%; Pred. No. 0.0084;
Matches 32; Conservative 20; Mismatches 57; Indels 26; Gaps 5;

QY 45 EPPPELENNK-----TNNRANGRPRHPFETKDYSEYSESCHELFHTRYVYDGCGRSAKRY 100
Db 129 EPPRKDAKKFMDHFWLNRNSAEEVLPRTKNEHDETCTLPFSQSVAHSECEKV-IV 187

QY 101 TELVCSGCGPARLLPNAIGRGKWMRPSGPDFR-----CIPDRYRAQRYOLLCPGGEA 153
Db 188 ONNLCEFGKCSS-----FHVPGPDRLTYTFCSCKLPYFMSKMFIDNCT-SSV 233

QY 154 PRAKRYLVASCKCK 168
Db 234 PVYKKWIVVECNCE 248

RESULT 4
Q9PUK2 PRELIMINARY; PRT: 272 AA.
ID Q9PUK2;
AC Q9PUK2;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
DE CARONTE.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Rodriguez Esteban C., Capdevilla J., Economides A.N., Pascual J.,
Ortiz A., Izpisua Belmonte J.C.;
RT "Caronte, a novel cer-lik protein, mediates the establishment of
embryonic left-right asymmetry."
RL Nature 0:0-0(1999).
DR EMBL: AF179484; AAD55581.1;
DR InterPro: IPR000359;
DR ProDom: PD000357;
DR PROSITE: PS01225; CTCK_2; 1.
DR SMART: SM00041; CT: 1.
SQ SEQUENCE 272 AA; 31201 MW; 0D89729715771BC2 CRC64;

Query Match 8.9%; Score 104; DB 13; Length 272;
Best Local Similarity 23.7%; Pred. No. 0.013;
Matches 32; Conservative 19; Mismatches 58; Indels 26; Gaps 5;

QY 45 EPPPELENNK-----TNNRANGRPRHPFETKDYSEYSESCHELFHTRYVYDGCGRSAKRY 100
Db 129 EPPRKDAKKFMDHFWLNRNSAEEVLPRTKNEHDETCTLPFSQSVAHSECEKV-IV 187

QY 101 TELVCSGCGPARLLPNAIGRGKWMRPSGPDFR-----CIPDRYRAQRYOLLCPGGEA 153
Db 188 ONNLCEFGKCSS-----FHVPGPDRLTYTFCSCKLPYFMSKMFIDNCT-SSV 233

QY 154 PRAKRYLVASCKCK 168
Db 234 PVYKKWIVVECNCE 248

RESULT 5
Q88273 PRELIMINARY; PRT: 168 AA.
ID Q88273;
AC Q88273;
DT 01-NOV-1998 (TREMblrel. 08, Created)
DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)

```

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2001, 15:48:41 ; Search time 80.28 Seconds  
(without alignments)  
160.848 Million cell updates/sec

Title: US-09-668-021-6

Perfect score: 1165

Sequence: 1 MQLPLALCLICLVHTAFRV.....KPRPRARSAKANQALEENAY 213

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Archived: 412676 seqs, 60623988 residues

412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 405 summaries

Database :

A\_Geneseq\_0601:\*

1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT:\*

2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT:\*

3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT:\*

4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT:\*

5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT:\*

6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT:\*

7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT:\*

8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT:\*

9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT:\*

10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT:\*

11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT:\*

12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT:\*

13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT:\*

14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT:\*

15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT:\*

16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT:\*

17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT:\*

18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT:\*

19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT:\*

20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT:\*

21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT:\*

22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1165	100.0	213	21	AAV96430	Human TGF-beta bin
2	1164	99.9	213	21	AAV96436	Human TGF-beta bin
3	1157	99.3	213	21	AAV96430	Human TGF-beta bin
4	1157	99.3	213	21	AAV96436	Human TGF-beta bin
5	1157	99.3	213	22	AAV97589	Human secreted pro
6	1136	97.5	213	21	AAV96431	Verret TGF-beta bi
7	1062	91.2	367	21	AAV96430	Human DAN/Cerberus
8	1040	89.3	213	21	AAV96433	Rat TGF-beta bindi
9	1021	87.6	211	21	AAV96432	Murine TGF-beta bi
10	892.5	76.6	176	21	AAV96434	Bovine TGF-beta bi
11	351.5	30.2	206	21	AAV75981	Murine skin cell p

12	351.5	30.2	206	21	AAV76031	Murine skin cell p
13	351.5	30.2	206	22	AAV55920	Skin cell protein,
14	351.5	30.2	206	22	AAV55970	Human small CCN-11
15	349.5	30.0	206	18	AAV09408	Human small CCN-11
16	349.5	30.0	206	19	AAV58704	Human adult retina
17	349.5	30.0	206	20	AAV95711	Human adult retina
18	349.5	30.0	206	21	AAV10233	Human 5' EST secre
19	256	22.0	50	20	AAV12009	Secreted protein A
20	122.5	10.5	116	18	AAV27654	Human secreted pro
21	122.5	10.5	116	18	AAV44090	Mutant human TGF-b
22	117	10.0	23	21	AAV96435	Amino acid sequenc
23	108	9.3	168	21	AAV84014	Human fetal kidney
24	104.5	9.0	102	20	AAV03225	Murine cerebral ne
25	104.5	9.0	102	21	AAV10277	Human b57 protein
26	103	8.8	168	21	AAV51132	Human b57 protein
27	92	7.9	184	19	AAV6293	Mouse gremlin. Ho
28	92	7.9	184	20	AAV42173	Human gremlin. Ho
29	92	7.9	184	21	AAV95961	Human DRM protein
30	90	7.7	184	20	AAV42172	Rat DRM protein se
31	90	7.7	184	20	AAV42174	EGFP/DRM fusion pr
32	90	7.7	429	20	AAV42175	DAN and b57 protei
33	87.5	7.5	184	19	AAV29727	Herpes simplex vir
34	87	7.5	246	21	AAV78333	T-PA with -ve char
35	87	7.5	391	20	AAV42178	DAN and b57 protei
36	87	7.5	558	12	AAV13150	T-PA with -ve char
37	86.5	7.4	182	19	AAV29726	Xenopus cerberus p
38	83.5	7.2	561	12	AAV12342	Human cancer assoc
39	83	7.1	270	19	AAV41250	Drosophila SLIT pr
40	82	7.0	150	21	AAV43879	Human tpa. Homo s
41	81.5	7.0	1480	13	AAV25079	Tissue plasminogen
42	81	7.0	562	6	AAV50342	Sequence of tissue
43	81	7.0	562	9	AAV82582	
44	81	7.0	562	11	AAV09288	
45	81	7.0	562	11	AAV09290	

#### ALIGNMENTS

##### RESULT 1

AAV96430 standard; Protein; 213 AA.

AAV96430:

12-SEP-2000 (first entry)

Human TGF-beta binding protein (BEER) variant V101.

osteopathic; transforming growth factor-beta; TGF-beta; binding protein;

BEER; variant; V101; gene therapy; antisense therapy; fracture;

chromosome 17q12-21; bone mineralization.

Homo sapiens.

Key Location/Qualifiers

Misc-difference 10 /label= V101

FT /note= "wild type valine has been substituted with

FT Isoleucine"

WO200032773-A1.

08-JUN-2000.

24-NOV-1999; 99WO-US27990.

27-NOV-1998; 98US-0110283.

(DARW-) DARWIN DISCOVERY LTD.

Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepier BW;

Van Ness J, Winkler DG;

XX MPI: 2000-412321/35.  
DR N-PSDB; AAA29056.  
XX Nucleic acids (1) encoding a transforming growth factor beta binding  
PT protein, useful for identifying agents for treating osteopenia,  
PT osteoporosis and fractures  
XX  
XX Claim 3; Page 119-120; 162pp; English.  
PS  
XX This shows a variant human transforming growth factor-beta (TGF-beta)  
CC binding protein designated BEER V101, which comprises a substitution of  
CC isoleucine for the wild-type valine at residue 10. The cDNA and protein  
CC may be used for prevention, treatment and diagnosis of diseases  
CC associated with inappropriate BEER expression. For example, they may be  
CC used to treat disorders associated with decreased TGF-beta BP expression.  
CC The cDNA or vectors may be administered to treat diseases by rectifying  
CC mutations or deletions in a patient's genome that affect the activity of  
CC BEER by expressing inactive proteins or to supplement the patients' own  
CC production of BEER polypeptides. The nucleic acids may be used for  
CC recombinant production of BEER, gene therapy, antisense therapy, as  
CC probes for diagnostic assays and for functional studies. BEER may be used  
CC to raise antibodies and for identification of BEER modulators. BEER  
CC antagonists may be used to increase bone mineral content for the  
CC treatment of disorders such as osteopenia, osteoporosis, fractures and  
CC other disorders associated with low mineral content.  
XX  
SQ Sequence 213 AA;  
  
Query Match 100.0%; Score 1165; DB 21; Length 213;  
Best Local Similarity 100.0%; Pred. No. 2.1e-105;  
Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MOLPLALCLTCLVHTAFRVESGQMGAFKNDATETIRLEGYEPPEPLENNKTNRAE 60  
DB 1 mqplalcltcllvhtafvveggwqafkndatellrelgeyeppeplennktnrae 60  
QY 61 NGRPRPHHFEETKDVSEYSCRELHFTRYVTDGPCRSAPVTELVCSGCGPARLLPNAIG 120  
DB 61 ngrpprhpfetkdvseyscrelhfttryvtdgpsraskpvtelvcsgcgparllpnalg 120  
QY 121 RCKWWRPSGDPDFRCIPDRRAQRVQLCPGEGAPRAKRVLVASCKKRLTRFHNOSELK 180  
DB 121 rkwmpsgdpdfrcipdrraqrvqlcpggeapraarkrvlvascckrltrfhngselk 180  
QY 181 DFGTEAARPOKGRKPRPARSAKANQAELENAY 213  
DB 181 dfgteaarpqkgrkprparsakangaelenay 213  
SQ  
RESULT 2  
ID AAY96436  
XX AAY96436 standard; Protein: 213 AA.  
AC AAY96436;  
XX  
DT 12-SEP-2000 (first entry)  
XX  
DE Human TGF-beta binding protein (BEER) variant P38R.  
XX  
KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;  
KW BEER; variant; P38R; gene therapy; antisense therapy; fracture;  
KW chromosome 17q12-21; bone mineralization.  
XX  
OS Homo sapiens.  
XX  
EH Key Location/Qualifiers  
FT Misc-difference 38 /Label= P38R  
FT /note= "wild type proline has been substituted with  
FT arginine"  
XX

PN WO200032773-A1.  
XX  
XX 08-JUN-2000.  
PD  
XX  
XX 24-NOV-1999; 99WO-US27990.  
PF  
XX  
XX 27-NOV-1998; 98US-0110283.  
PR  
XX  
XX (DARW-) DARWIN DISCOVERY LTD.  
PA  
XX  
XX Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paeper BW;  
PI Van Ness J, Winkler DG;  
PT N-PSDB; AAA29062.  
DR  
XX MPI: 2000-412321/35.  
DR  
XX Nucleic acids (1) encoding a transforming growth factor beta binding  
PT protein, useful for identifying agents for treating osteopenia,  
PT osteoporosis and fractures  
XX  
XX Disclosure; Page 121; 162pp; English.  
PS  
XX This shows a variant human transforming growth factor-beta  
CC (TGF-beta) binding protein designated BEER P38R. The encoded protein  
CC comprises a substitution of arginine for the wild-type proline at  
CC residue 38. The cDNA and protein may be used for prevention, treatment  
CC and diagnosis of diseases associated with inappropriate BEER expression.  
CC For example, they may be used to treat disorders associated with  
CC decreased TGF-beta BP expression. The cDNA or vectors may be administered  
CC to treat diseases by rectifying mutations or deletions in a patient's  
CC genome that affect the activity of BEER by expressing inactive proteins  
CC or to supplement the patients' own production of BEER polypeptides. The  
CC nucleic acids may be used for recombinant production of BEER, gene  
CC therapy, antisense therapy, as probes for diagnostic assays and for  
CC functional studies. BEER may be used to raise antibodies and for  
CC identification of BEER modulators. BEER antagonists may be used to  
CC increase bone mineral content for the treatment of disorders such as  
CC osteopenia, osteoporosis, fractures and other disorders associated with  
CC low mineral content.  
XX  
SQ Sequence 213 AA;  
  
Query Match 99.9%; Score 1164; DB 21; Length 213;  
Best Local Similarity 99.5%; Pred. No. 2.7e-105;  
Matches 212; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MOLPLALCLTCLVHTAFRVESGQMGAFKNDATETIRLEGYEPPEPLENNKTNRAE 60  
DB 1 mqplalcltcllvhtafvveggwqafkndatellrelgeyeppeplennktnrae 60  
QY 61 NGRPRPHHFEETKDVSEYSCRELHFTRYVTDGPCRSAPVTELVCSGCGPARLLPNAIG 120  
DB 61 ngrpprhpfetkdvseyscrelhfttryvtdgpsraskpvtelvcsgcgparllpnalg 120  
QY 121 RCKWWRPSGDPDFRCIPDRRAQRVQLCPGEGAPRAKRVLVASCKKRLTRFHNOSELK 180  
DB 121 rkwmpsgdpdfrcipdrraqrvqlcpggeapraarkrvlvascckrltrfhngselk 180  
QY 181 DFGTEAARPOKGRKPRPARSAKANQAELENAY 213  
DB 181 dfgteaarpqkgrkprparsakangaelenay 213  
SQ  
RESULT 3  
ID AAB26106  
XX AAB26106 standard; Protein: 213 AA.  
AC AAB26106;  
XX  
DT 15-JAN-2001 (first entry)  
XX  
DE Human DAN/Cerberus-related protein 6 (hDNR6) (exons 1 and 4).



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2001, 15:47:15 : Search time 47.86 Seconds  
(without alignments)  
100.150 Million cell updates/sec

Title: US-09-668-021-6

Perfect score: 1165

Sequence: 1 MDPLALCLICLVHFAFRV.....KPRPRASAKAQAELFNAY 213

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

arched: 212252 segs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata2/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata2/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata2/1aa/5A\_COMB.pep:\*  
4: /cgn2\_6/ptodata2/1aa/5B\_COMB.pep:\*  
5: /cgn2\_6/ptodata2/1aa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata2/1aa/backfill1st.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	351.5	30.2	206	4	US-09-188-930-159 Sequence 159, App
2	351.5	30.2	206	4	US-09-188-930-286 Sequence 286, App
3	349.5	30.0	206	1	US-08-468-847B-2 Sequence 2, Appli
4	349.5	30.0	206	1	US-08-468-847B-20 Sequence 20, Appli
5	87.5	7.5	735	3	US-09-191-647-9 Sequence 9, Appli
6	87.5	7.5	735	4	US-09-540-245A-9 Sequence 9, Appli
7	87.5	7.5	735	4	US-09-540-153-9 Sequence 9, Appli
8	82	7.0	270	4	US-08-878-474-1 Sequence 1, Appli
9	81.5	7.0	1480	3	US-09-191-647-7 Sequence 7, Appli
10	81.5	7.0	1480	4	US-09-540-245A-7 Sequence 7, Appli
11	81.5	7.0	1480	4	US-09-540-153-7 Sequence 7, Appli
12	81.5	7.0	1480	5	PCT-US91-09055-2 Sequence 2, Appli
13	80	6.9	510	4	US-08-246-489-2 Sequence 2, Appli
14	79	6.8	336	4	US-08-804-166-8 Sequence 8, Appli
15	79	6.8	336	4	US-08-910-991-8 Sequence 8, Appli
16	79	6.8	566	1	US-08-073-383-4 Sequence 4, Appli
17	79	6.8	566	5	PCT-US94-06365-4 Sequence 4, Appli
18	78	6.7	1210	2	US-08-484-438-7 Sequence 7, Appli
19	78	6.7	1210	2	US-08-475-035-4 Sequence 4, Appli
20	77.5	6.7	75	5	PCT-US91-09055-9 Sequence 9, Appli
21	76	6.5	546	6	5200340-6 Patent No. 5200340
22	76	6.5	562	2	US-08-560-098A-50 Sequence 50, Appli
23	76	6.5	562	2	US-08-883-795A-38 Sequence 38, Appli
24	76	6.5	562	6	5185259-3 Patent No. 5185259
25	76	6.5	562	6	5200340-2 Patent No. 5200340
26	76	6.5	562	6	5344773-2 Patent No. 5344773
27	74	6.4	470	2	US-08-946-241B-2 Sequence 2, Appli

28	74	6.4	470	3	US-09-309-053-2 Sequence 2, Appli
29	74	6.4	479	2	US-08-946-241B-9 Sequence 9, Appli
30	74	6.4	479	3	US-09-309-053-9 Sequence 9, Appli
31	73.5	6.3	301	4	US-08-303-861-21 Sequence 21, Appli
32	73.5	6.3	301	4	US-09-011-073A-1 Sequence 1, Appli
33	73.5	6.3	685	2	US-08-878-989-1 Sequence 1, Appli
34	73.5	6.3	685	2	US-09-136-282-2 Sequence 2, Appli
35	73.5	6.3	685	4	US-09-272-796-1 Sequence 1, Appli
36	73.5	6.3	685	4	US-09-505-744-2 Sequence 2, Appli
37	73	6.2	301	4	US-09-230-421-2 Sequence 2, Appli
38	72.5	6.2	375	1	US-08-468-847B-13 Sequence 13, Appli
39	72	6.2	379	1	US-08-468-847B-11 Sequence 11, Appli
40	71.5	6.1	521	3	US-08-948-564-14 Sequence 14, Appli
41	71	6.1	449	1	US-08-102-942A-4 Sequence 4, Appli
42	71	6.1	1259	3	US-09-187-049-13 Sequence 13, Appli
43	70.5	6.1	390	1	US-08-347-792-15 Sequence 15, Appli
44	70.5	6.1	390	1	US-08-431-357-15 Sequence 15, Appli
45	70.5	6.1	390	4	US-08-392-542-3 Sequence 3, Appli

## ALIGNMENTS

```
RESULT 1
US-09-188-930-159
; Sequence 159, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Marison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 159
; LENGTH: 206
; TYPE: PRT
; ORGANISM: mouse
US-09-188-930-159

Query Match          30.2%   Score 351.5;   DB 4;   Length 206;
Best Local Similarity 40.4%;   Pred. No. 5.9e-30;
Matches 82;   Conservative 32;   Mismatches 68;   Indels 21;   Gaps 8;

QY 9 LIGLVHTAFRRVGGQWQAFKNDATETIRLGEYPEPPPELNNKTMRAENGRRPHH 68
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 13 LILCILRNCL-----AFKNDATETILYSHVVKP-VPAPHSNSTLNQARNGCR--HF 60

QY 69 PFETKIVS---EYSCRELHFPRYVDSGCRSAKPVTELNSGCGPARLPAIAGSG--- 122
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 SSTGLDRNRSVVQCGRELSTKSTISDGCTISPLKELVAGBCLPLVPLNNIGGYGT 120

QY 123 KMW-RPSGDFPCIDPDRYRAQRVQLCPGGEAPARARVRLVASCCKRLRFRNOSLKD 181
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 121 KYWRSRSGEMCVNDKPTQRIQLQCCDG--STRTYKLTVTACKCRKYTRQNNESHNF 179

QY 182 FCTEARAPQKGRKPRPRA-RSAK 203
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 180 ESYVPAKPAQHRRKRRASKSK 202

RESULT 2
US-09-188-930-286
; Sequence 286, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
```

APPLICANT: Watson, James D.  
APPLICANT: Strachan, Lorna  
APPLICANT: Sleeman, Matthew  
APPLICANT: Murrison, Rene  
APPLICANT: Murrison, James Greg  
TITLE OF INVENTION: Compositions Isolated From Skin Cells  
FILE REFERENCE: 11000.1011C1  
CURRENT APPLICATION NUMBER: US/09/188,930A  
CURRENT FILING DATE: 1998-11-09  
NUMBER OF SEQ ID NOS: 348  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 286  
LENGTH: 206  
TYPE: PR1  
ORGANISM: Mouse  
US-09-188-930-266

Query Match 30.2%; Score 351.5; DB 4; Length 206;  
Best Local Similarity 40.4%; Pred. No. 5.9e-30;  
Matches 82; Conservative 32; Mismatches 68; Indels 21; Gaps 8;

9 LCLILVAFRYVEGOGWAFKNDATETIRELGEYPPPELENNKTWNRANGRPPHH 68  
13 LCLILVAFRYVEGOGWAFKNDATETIRELGEYPPPELENNKTWNRANGRPPHH 68  
69 PFETKDVSS--EYSCRELHFTRYVTDGPCRSAPVTELVCSGCGPARLLPNAI 122  
61 SSTGLRNSRVSQWRCVNDKRTORIOLQCGDC-STRTYKITYVTACKCKRYTQHSHSNF 120  
123 KMW-RPSGDPFRCIPDRYRAORVQLCPGGEAPRARVRLVASCCKRLTFPHNSELKD 181  
121 KYMSRRSSQWRCVNDKRTORIOLQCGDC-STRTYKITYVTACKCKRYTQHSHSNF 179  
182 FGTBARPOKGRKPPRA-RSAK 203  
180 ESVSPAKPAQHRRKRASKSSK 202

RESULT 3  
US-08-468-847B-2  
Sequence 2, Application US/08468847B  
Patent No. 5780263  
GENERAL INFORMATION:  
APPLICANT: Hastings, Gregg A. and Adams, Mark D.  
TITLE OF INVENTION: Human CCN-Like Growth Factor  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,  
ADDRESSEE: CECCHI, STEWART & OLSTEIN  
STREET: 6 BECKER FARM ROAD  
CITY: ROSELAND  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 INCH DISKETTE  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/468,847B  
FILING DATE: 6 June 1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: MULLINS, J.G.  
REGISTRATION NUMBER: 33,073  
REFERENCE/DOCKET NUMBER: 325800-442  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 206 AMINO ACIDS  
TYPE: AMINO ACID  
STRANDEDNESS:  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PROTEIN  
US-08-468-847B-2

Query Match 30.0%; Score 349.5; DB 1; Length 206;  
Best Local Similarity 40.7%; Pred. No. 9.7e-30;  
Matches 85; Conservative 32; Mismatches 67; Indels 25; Gaps 9;

3 LPLALICLILVAFRYVEGOGWAFKNDATETIRELGEYPPPELENNKTWNRANG 62  
11 LPLA-----CLIKSCL-----AFKNDATETILSHVVKP-VPAHPSNSTLNQARNG 56  
63 GRPHHPETKDVSS--EYSCRELHFTRYVTDGPCRSAPVTELVCSGCGPARLLPNAI 119  
57 GR--HFSNTGLDRMTRYOVGCRELRSTKYISDQCTSIPLKELVACAGECPLPVLPMWI 114  
120 GRG--KMW-RPSGDPFRCIPDRYRAORVQLCPGGEAPRARVRLVASCCKRLTFPHN 175  
115 GGGYGTKYMSRRSSQWRCVNDKRTORIOLQCGDC-STRTYKITYVTACKCKRYTQHSH 173  
176 OSELKDFGTBARPOKGRKPPRA-RSAK 203  
174 ESHNFSMSPAKPAQHRRKRASKSSK 202

RESULT 4  
US-08-468-847B-20  
Sequence 20, Application US/08468847B  
Patent No. 5780263  
GENERAL INFORMATION:  
APPLICANT: Hastings, Gregg A. and Adams, Mark D.  
TITLE OF INVENTION: Human CCN-Like Growth Factor  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,  
ADDRESSEE: CECCHI, STEWART & OLSTEIN  
STREET: 6 BECKER FARM ROAD  
CITY: ROSELAND  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 INCH DISKETTE  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/468,847B  
FILING DATE: 6 June 1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: MULLINS, J.G.  
REGISTRATION NUMBER: 33,073  
REFERENCE/DOCKET NUMBER: 325800-442  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 206 AMINO ACIDS  
TYPE: AMINO ACID  
STRANDEDNESS:











```

RESULT 2
ID 09PWB0 PRELIMINARY; PRT; 272 AA.
AC 09PWB0;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE CERBERUS HOMOLOG.
GN CER.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99439862; Pubmed=10508582;
RA Zhu L., Marvin M.J., Gardiner A., Lassar A.B., Mercola M., Stern C.D.,
RA Levin M.;
RT "Cerberus regulates left-right asymmetry of the embryonic head and
heart."
Curr. Biol. 9:931-938(1999).
EMBL: AF139721; AAD51610.1; -.
InterPro: IPR000359; -.
DR PROSITE: PS01225; CTKC_2; 1.
DR SMART: SM00041; CT; 1.
SQ SEQUENCE 272 AA; 31224 MW; 36E4C9F719711BCA CRC64;

Query Match 9.2%; Score 107; DB 13; Length 272;
Best Local Similarity 23.7%; Pred. No. 0.0064;
Matches 32; Conservative 21; Mismatches 56; Indels 26; Gaps 5;

QY 45 EPPPELENNK-----TNNRAENGRRPHHPETKDVSEYSCRELHFRVYTDGPCRSAPV 100
DB 129 EPPYRKDAKKFMDHFMRLKNSASEEVLDPIKTNEMHQETCRLLPFQSVAHNSCEKV-IV 187
QY 101 TELVSCGCGPARLLPNAIGRKWMRPSGPDFR-----CIPDRYRAQVOLLCPGGA 153
DB 188 QNNLCFGKCSS-----FHVPGPDRLYTFCKCLPTFKFSKMHDLNCT-SSV 233
QY 154 PRAKRYRLVASCKCK 168
DB 234 PYYKRWIVVEECNCE 248

RESULT 3
ID 09PUK2 PRELIMINARY; PRT; 272 AA.
AC 09PUK2;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE CARONTE.
GN CAR.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX Rodriguez Esteban C., Capdevilla J., Economides A.N., Pascual J.,
RA Ortiz A., Izpisua Belmonte J.C.;
RT "Caronte, a novel cer-like protein, mediates the establishment of
embryonic left-right asymmetry."
RL Nature 0:0-0(1999).
EMBL: AF179484; AAD5581.1; -.
InterPro: IPR000359; -.
DR InterPro: IPR001839; -.
DR ProDom: PD000357; -.
DR PROSITE: PS01225; CTKC_2; 1.

```

```

DR SMART: SM00041; CT; 1.
SQ SEQUENCE 272 AA; 31201 MW; 0D89729715771BC2 CRC64;

Query Match 9.0%; Score 105; DB 13; Length 272;
Best Local Similarity 23.7%; Pred. No. 0.01;
Matches 32; Conservative 20; Mismatches 57; Indels 26; Gaps 5;

QY 45 EPPPELENNK-----TNNRAENGRRPHHPETKDVSEYSCRELHFRVYTDGPCRSAPV 100
DB 129 EPPYRKDAKKFMDHFMRLKNSASEEVLDPIKTNEMHQETCRLLPFQSVAHNSCEKV-IV 187
QY 101 TELVSCGCGPARLLPNAIGRKWMRPSGPDFR-----CIPDRYRAQVOLLCPGGA 153
DB 188 QNNLCFGKCSS-----FHVPGPDRLYTFCKCLPTFKFSKMHDLNCT-SSV 233
QY 154 PRAKRYRLVASCKCK 168
DB 234 PYYKRWIVVEECNCE 248

RESULT 4
ID 09H772 PRELIMINARY; PRT; 168 AA.
AC 09H772;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE CDNA: FLJ21195 FIS, CLONE COL00185.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLON;
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RT "NEDD human cDNA sequencing project."
RL Submitted (AUG-2000) to the EMBL/Genbank/DBD databases.
EMBL: AK024648; BAB15026.1; -.
DR AK024648; BAB15026.1; -.
SQ SEQUENCE 168 AA; 19320 MW; D5A4E4E818BF8C0E CRC64;

Query Match 8.5%; Score 99; DB 4; Length 168;
Best Local Similarity 24.5%; Pred. No. 0.025;
Matches 45; Conservative 25; Mismatches 66; Indels 48; Gaps 9;

QY 2 QLPALCLVCLLVHAAFRVVEGGQWQAFKNDATETIIPELGEYPEPPPELENNKTNRAEN 61
DB 4 KLSLSTFLVAVLVKVA-----EARKNRPA-----GALSPYKDGSSN-----N 41
QY 62 GGRPHHPETKDVSEYSCRELHFRVYTDGPCRSAPV 100
DB 42 SERMOHQIKFVLSQDALVYTERKYLKSDCKTOPLRVSEEGRS-RTLNRFCTGG 100
QY 109 CGPARLLPNAIGRKWMRPSGPDFR-----CIPDRYRAQVOLLCPGGAAP-PARKRVRLVA 163
DB 101 CN-SFYPRHV-----KKEESFQSCAFCKPQRYVSVLELECGDLPDPFLKKIQKVK 153
QY 164 SKC 167
DB 154 QCRK 157

RESULT 5
ID 09UEM9 PRELIMINARY; PRT; 493 AA.
AC 09UEM9;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)

```



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2001, 15:48:42 ; Search time 80.28 Seconds  
(without alignments)  
160.848 Million cell updates/sec

Title: US-09-668-021-10

Perfect score: 1169

Sequence: 1 MQLPLALCLVCLVHNAFRV.....KPRPRARAKANQAELENAY 213

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Residues: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_0601.\*

- 1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT.\*
- 2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT.\*
- 3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT.\*
- 4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT.\*
- 5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT.\*
- 6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT.\*
- 7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT.\*
- 8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT.\*
- 9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT.\*
- 10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT.\*
- 11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT.\*
- 12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT.\*
- 13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT.\*
- 14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT.\*
- 15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT.\*
- 16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT.\*
- 17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT.\*
- 18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT.\*
- 19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT.\*
- 20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT.\*
- 21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT.\*
- 22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1169	100.0	213	21	AAV96431
2	1146	98.0	213	21	AAAB26106
3	1146	98.0	213	21	AAV96429
4	1146	98.0	213	22	AAV97589
5	1137	97.3	213	21	AAV96436
6	1136	97.2	213	21	AAV96430
7	1073	90.8	213	21	AAV96433
8	1052	90.0	211	21	AAV96432
9	1051	89.9	367	21	AAAB26105
10	920.5	78.7	176	21	AAV96434
11	348	29.8	206	18	AAW09408

12	348	29.8	206	19	AAW58704
13	348	29.8	206	20	AAW95711
14	348	29.8	206	21	AAAB10233
15	346	29.6	206	21	AAV75981
16	346	29.6	206	21	AAV76031
17	346	29.6	206	22	AAW55920
18	346	29.6	206	22	AAW55970
19	261	22.3	50	20	AAV12009
20	122.5	10.5	116	18	AAW27654
21	122.5	10.5	116	18	AAW44090
22	113	9.7	23	21	AAV96435
23	103.5	8.9	102	21	AAV03225
24	103.5	8.9	102	21	AAV10277
25	99	8.5	168	21	AAV84014
26	94	8.0	168	21	AAV51132
27	93	8.0	712	21	AAW45147
28	92	7.9	558	12	AAV13150
29	89.5	7.7	1428	21	AAV97033
30	89	7.6	336	18	AAW33360
31	88.5	7.6	561	12	AAV12342
32	88	7.5	184	19	AAW69293
33	88	7.5	184	20	AAV42173
34	88	7.5	184	21	AAV95961
35	86.5	7.4	870	21	AAW41742
36	86	7.4	562	6	AAV50342
37	86	7.4	562	9	AAW82582
38	86	7.4	562	11	AAW09288
39	86	7.4	562	11	AAW09290
40	85.5	7.3	145	12	AAV15109
41	85.5	7.3	145	12	AAV15125
42	85	7.3	562	11	AAW09289
43	84.5	7.2	184	19	AAW29727
44	84.5	7.2	184	20	AAV42172
45	84.5	7.2	184	20	AAV42174

#### ALIGNMENTS

RESULT 1					
ID	AAV96431	standard:	Protein:	213	AA.
XX	XX				
AC	AAV96431:				
XX	XX				
DT	12-SEP-2000	(first entry)			
XX	XX				
DE	Vervet TGF-beta binding protein (BEER).				
KW	osteopathic; transforming growth factor-beta; TGF-beta; binding protein;				
KM	BEER; gene therapy; antisense therapy; fracture; bone mineralization.				
XX	XX				
OS	Cercopithecus pygerythrus.				
XX	XX				
PN	WO200032773-A1.				
XX	XX				
PD	08-JUN-2000.				
XX	XX				
PP	24-NOV-1999:	99WO-US27990.			
XX	XX				
PR	27-NOV-1998:	98US-0110283.			
XX	XX				
PA	(DARW-) DARWIN DISCOVERY LTD.				
XX	XX				
PI	Brunow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepers BW;				
XX	Van Ness J, Winkler DG;				
XX	XX				
DR	WPI; 2000-412321/35.				
XX	XX				
PT	Nucleic acids (I) encoding a transforming growth factor beta binding				
XX	protein, useful for identifying agents for treating osteopenia,				
PT	osteoporosis and fractures				

Human small CCN-11  
Homo sapiens fetal  
Human adult retina  
Murine skin cell p  
Murine skin cell p  
Skin cell protein,  
Skin cell protein,  
Human 5' EST secre  
Secreted protein A  
Human secreted pro  
Mutant human TGF-  
Amino acid sequenc  
Human fetal kidney  
Amino acid sequenc  
Murine cerebral ne  
Human ORFX ORF2911  
T-PA with -ve char  
Caspase 8-interact  
TBP(20-190)/hcg-be  
T-PA with -ve char  
Human b57 protein  
Mouse DRM protein  
Human gremilin. Ho  
Human ORFX ORF1506  
Human tpa. Homo s  
Tissue plasminogen  
Sequence of tissue  
hcg/elh chimera, E  
hcg/HLH chimera, A  
Sequence of tissue  
DAN and b57 protei  
Human DRM protein se  
Rat DRM protein se

```
XX Claim 4; Page 122-123; 162pp; English.
PS This shows a veyret transforming growth factor-beta (TGF-beta)
CC binding protein designated vBEER. The cDNA and protein may be used for
CC prevention, treatment and diagnosis of diseases associated with
CC inappropriate BEER expression. For example, they may be used to treat
CC disorders associated with decreased TGF-beta BP expression. The cDNA or
CC vectors may be administered to treat diseases by rectifying mutations or
CC deletions in a patient's genome that affect the activity of BEER by
CC expressing inactive proteins or to supplement the patient's own production
CC of BEER polypeptides. The nucleic acids may be used for recombinant
CC production of BEER, gene therapy, antisense therapy, as probes for
CC diagnostic assays and for functional studies. BEER may be used to raise
CC antibodies and for identification of BEER modulators. BEER antagonists
CC may be used to increase bone mineral content for the treatment of
CC disorders such as osteopenia, osteoporosis, fractures and other disorders
CC associated with low mineral content.
XX
SQ Sequence 213 AA;

Query Match 100.0%; Score 1169; DB 21; Length 213;
Best Local Similarity 100.0%; Pred. No. 9e-106;
Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQLPLALCLVCLLHAAFRVVGSGQMAFKNDATETIPELGEPPEPPPELNNKTMNRAE 60
Db 1 mqlplalclvcllvhaafrrvvgsgwqafkndateilpelgpeppeppelennktnmrae 60
QY 61 NCGRRPHHPETKDVSEYSCRELHFTRYVDGPCRSAPVTVELVCSGCGPARLLPNAIG 120
Db 61 ngrrpnhphtekdvseyscrelhfttryvtdgpcrsakpvtelvcsgcgparllpnaig 120
QY 121 RKGMMWRSPGDFPCIPDRYRAORVOLLCPGGAAPRAKRVRLVASCKCKRLTFHNOSELK 180
Db 121 rkgmmwrspgdfpcipdryraqvqlldpggaapraarkrvrlvasckckrltrfhngselk 180
QY 181 DFGPEAARPOKGRKPRPARAKAQAQAELENAY 213
Db 181 dfgpeaarpqkgrkprparargakangaelenay 213

RESULT 2
AAB26106
ID AAB26106 standard; Protein: 213 AA.
XX
AC AAB26106;
XX
XX 15-JAN-2001 (first entry)
XX
DE Human DAN/Cerberus-related protein 6 (hDCR6) (exons 1 and 4).
XX
KW Human; DNA/Cerberus-related protein 6; hDCR6; morphogenic protein;
KW antagonist; BMP; cell growth; cell differentiation; bone formation;
KW gene therapy.
XX
OS Homo sapiens.
XX
XX WO200055193-A2.
XX
XX 21-SEP-2000.
XX
XX 02-MAR-2000; 2000WO-US05537.
XX
XX 12-MAR-1999; 99US-0124118.
XX
XX (REGF-) REGENERON PHARM INC.
XX
XX Economides AN;
XX
XX WPI; 2000-638179/61.
XX
XX N-PSDB; AAA94051.
XX
DR
```

```
XX Novel isolated, human DNA/Cerberus related protein 6 which include
PT natural homologue, and polypeptides comprising DCR6 domain and nucleic
PT acids encoding the proteins which are useful as probes and primers
XX
XX Claim 8; Fig 3; 40pp; English.
XX
CC The present sequence comprises the amino acid sequence encoded by exons 1
CC and 4 of the human DAN/cerberus-related protein 6 (hDCR6) coding
CC sequence. The coding sequence was isolated from a human kidney cDNA
CC library containing exons 1 and 4 of the sequence. hDCR6 is closely
CC related to the DAN and DCR5 proteins, both of which act as antagonists of
CC morphogenic proteins such as BMP. It is possible that the hDCR6 gene and
CC protein can be used as immunogens, modulators of cell function, growth
CC and differentiation, to reduce undesirable bone formation, to identify
CC DCR6 binding agents, in diagnosis, and in gene therapy.
XX
SQ Sequence 213 AA;

Query Match 98.0%; Score 1146; DB 21; Length 213;
Best Local Similarity 98.1%; Pred. No. 1.5e-103;
Matches 209; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MQLPLALCLVCLLHAAFRVVGSGQMAFKNDATETIPELGEPPEPPPELNNKTMNRAE 60
Db 1 mqlplalclvcllvhaafrrvvgsgwqafkndateilpelgpeppeppelennktnmrae 60
QY 61 NCGRRPHHPETKDVSEYSCRELHFTRYVDGPCRSAPVTVELVCSGCGPARLLPNAIG 120
Db 61 ngrrpnhphtekdvseyscrelhfttryvtdgpcrsakpvtelvcsgcgparllpnaig 120
QY 121 RKGMMWRSPGDFPCIPDRYRAORVOLLCPGGAAPRAKRVRLVASCKCKRLTFHNOSELK 180
Db 121 rkgmmwrspgdfpcipdryraqvqlldpggaapraarkrvrlvasckckrltrfhngselk 180
QY 181 DFGPEAARPOKGRKPRPARAKAQAQAELENAY 213
Db 181 dfgleaarpqkgrkprparsakangaelenay 213

RESULT 3
AAY96429
ID AAY96429 standard; Protein: 213 AA.
XX
AC AAY96429;
XX
XX 12-SEP-2000 (first entry)
XX
DE Human TGF-beta binding protein (BEER).
XX
DE osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
KW BEER; gene therapy; antisense therapy; fracture; bone mineralization.
XX
OS Homo sapiens.
XX
XX WO200032773-A1.
XX
XX 08-JUN-2000.
XX
XX 24-NOV-1999; 99WO-US27990.
XX
XX 27-NOV-1998; 98US-0110283.
XX
XX (DARW-) DARWIN DISCOVERY LTD.
XX
XX Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepker BW;
XX Van Ness J, Winkler DG;
XX
XX WPI; 2000-412321/35.
XX
XX N-PSDB; AAA29055.
XX
XX Nucleic acids (I) encoding a transforming growth factor beta binding
PT
```

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2001, 15:47:16 ; Search time 47.86 Seconds  
(without alignments)  
100.150 Million cell updates/sec

Title: US-09-668-021-10

Perfect score: 1169

Sequence: 1 MQLPLALCLVCLLVHAAFRV.....KPRPRARAKANQAELENNAY 213

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5  
rched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/Backfilest1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	348	29.8	206	1	US-08-468-847B-2
2	348	29.8	206	1	US-08-468-847B-20
3	346	29.6	206	4	US-09-188-930-159
4	346	29.6	206	4	US-09-188-930-286
5	89	7.6	336	4	US-08-804-166-8
6	89	7.6	336	4	US-08-804-166-8
7	88.5	7.6	735	3	US-09-191-647-9
8	88.5	7.6	735	4	US-09-540-245A-9
9	88.5	7.6	735	4	US-09-540-153-9
10	81.5	7.0	1480	3	US-09-191-647-7
11	81.5	7.0	1480	4	US-09-540-245A-7
12	81.5	7.0	1480	4	US-09-540-153-7
13	81.5	7.0	1480	5	PCT-US91-09055-2
14	81	6.9	545	6	5200340-6
15	81	6.9	562	2	US-08-560-098A-50
16	81	6.9	562	2	US-08-883-795A-38
17	81	6.9	562	6	5185259-3
18	81	6.9	562	6	5200340-2
19	81	6.9	562	6	5344773-2
20	80	6.8	307	4	US-08-804-166-4
21	80	6.8	307	4	US-08-910-991-4
22	79.5	6.8	145	1	US-08-298-189B-1
23	79.5	6.8	145	1	US-08-475-213-10
24	79.5	6.8	145	1	US-08-395-238-2
25	79.5	6.8	145	4	US-09-142-320-12
26	79.5	6.8	145	4	US-09-142-320-13
27	79.5	6.8	145	4	US-09-142-320-14

28	79.5	6.8	145	4	US-09-142-320-15	Sequence 15, Appl
29	79.5	6.8	145	4	US-08-918-288-68	Sequence 68, Appl
30	79.5	6.8	145	4	US-09-282-357-68	Sequence 68, Appl
31	79.5	6.8	265	4	US-08-918-288-3	Sequence 3, Appl
32	79.5	6.8	265	4	US-08-918-288-39	Sequence 39, Appl
33	79.5	6.8	265	4	US-09-282-357-3	Sequence 3, Appl
34	79.5	6.8	265	4	US-09-282-357-39	Sequence 39, Appl
35	79	6.8	270	2	US-08-878-474-1	Sequence 2, Appl
36	79	6.8	470	2	US-08-946-241B-2	Sequence 2, Appl
37	79	6.8	470	3	US-09-309-053-2	Sequence 9, Appl
38	79	6.8	479	2	US-08-946-241B-9	Sequence 9, Appl
39	79	6.8	479	3	US-09-309-053-9	Sequence 9, Appl
40	78.5	6.7	165	2	US-08-709-924-2	Sequence 2, Appl
41	78.5	6.7	165	2	US-08-709-925-2	Sequence 2, Appl
42	78.5	6.7	181	4	US-08-918-288-36	Sequence 36, Appl
43	78.5	6.7	181	4	US-09-282-357-36	Sequence 36, Appl
44	78	6.7	510	4	US-08-246-489-2	Sequence 2, Appl
45	77.5	6.6	145	1	US-08-425-673-1	Sequence 1, Appl

## ALIGNMENTS

RESULT 1  
US-08-468-847B-2  
Sequence 2, Application US/08468847B  
Patent No. 5780263  
GENERAL INFORMATION:  
APPLICANT: Hastings, Gregg A. and Adams, Mark D.  
TITLE OF INVENTION: Human CCN-like Growth Factor  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,  
ADDRESSEE: CECCHI, STEWART & OLSTEIN  
STREET: 6 BECKER FARM ROAD  
CITY: ROSELAND  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 INCH DISKETTE  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/468,847B  
FILING DATE: 6 June 1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: MULLINS, J.G.  
REGISTRATION NUMBER: 33,073  
REFERENCE/DOCKET NUMBER: 325800-442  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ. ID NO. 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 206 AMINO ACIDS  
TYPE: AMINO ACID  
STRANDEDNESS:  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PROTEIN  
US-08-468-847B-2

Query Match 29.8%; Score 348; DB 1; Length 206;  
Best Local Similarity 40.5%; Pred. No. 3.1e-29;  
Matches 83; Conservative 31; Mismatches 65; Indels 26; Gaps 9;  
QY 3 LPALCLVCLLVHAAFRVVGQGMFKNDATETIIFELGEYEP-PPELENNKTNRAEN 61

```

Db 11 LPLA-----CILMKSCU-----AFKNDATEIL--YSHVVKPVPAHSSNSTLNOARN 55
Qy 62 GGRPHHPETKDV---EYSCRELHFTRYVTDGPCRSAPVTELYVCSGCGPARLLPNA 118
Db 56 GGR--HFSNTGLDRNTRVOYGCRLSTKTYISDSGCTISPLKELVCAECPLPLPVLPMW 113
Qy 119 IGRG---KWM-RPSGDPFCIPDRYRAORVOLLCPGGAAPRAKRVLVASCCKRLTRFH 174
Db 114 IGGYGTKTYMSRRSSQEMRCVNDKTRTORIQLOCDGST-RTYKITVYVACKCKRYTROH 172
Qy 175 NOSELKDFGEAPARPOKGRKPRPA 199
Db 173 NESHNFESMSPAKPVQHHRRKRA 197

```

RESULT 2  
US-08-468-847B-20

; Sequence 20, Application US/08468847B  
; Patent No. 5780263

; GENERAL INFORMATION:

; APPLICANT: Hastings, Gregg A. and Adams, Mark D.  
; TITLE OF INVENTION: Human CCN-Like Growth Factor  
; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,  
; ADDRESSEE: CECCHI, STEWART & OLSTEIN  
; STREET: 6 BECKER FARM ROAD  
; CITY: ROSELAND  
; STATE: NEW JERSEY  
; COUNTRY: USA  
; ZIP: 07068

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 INCH DISKETTE  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WORD PERFECT 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/468,847B  
; FILING DATE: 6 June 1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MULLINS, J G.  
; REGISTRATION NUMBER: 33,073  
; REFERENCE/DOCKET NUMBER: 325800-442  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-994-1700  
; TELEFAX: 201-994-1744  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 206 AMINO ACIDS  
; TYPE: AMINO ACID  
; STRANDEDNESS:  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: PROTEIN  
; US-08-468-847B-20

Query Match 29.8%; Score 348; DB 1; Length 206;  
Best Local Similarity 40.5%; Pred. No. 3.1e-29;  
Matches 83; Conservative 31; Mismatches 65; Indels 26; Gaps 9;

```

Qy 3 LPLALVCLVLAHFAFRVVGQGMQAFKNDATFIIPELGEYEP-PPLELNKKTMMRAEN 61
Db 11 LPLA-----CILMKSCU-----AFKNDATEIL--YSHVVKPVPAHSSNSTLNOARN 55
Qy 62 GGRPHHPETKDV---EYSCRELHFTRYVTDGPCRSAPVTELYVCSGCGPARLLPNA 118
Db 56 GGR--HFSNTGLDRNTRVOYGCRLSTKTYISDSGCTISPLKELVCAECPLPLPVLPMW 113

```

```

Qy 119 IGRG---KWM-RPSGDPFCIPDRYRAORVOLLCPGGAAPRAKRVLVASCCKRLTRFH 174
Db 114 IGGYGTKTYMSRRSSQEMRCVNDKTRTORIQLOCDGST-RTYKITVYVACKCKRYTROH 172
Qy 175 NOSELKDFGEAPARPOKGRKPRPA 199
Db 173 NESHNFESMSPAKPVQHHRRKRA 197

```

RESULT 3  
US-09-188-930-159

; Sequence 159, Application US/09188930A  
; Patent No. 6150502

; GENERAL INFORMATION:

; APPLICANT: Watson, James D.  
; APPLICANT: Strachan, Lorna  
; APPLICANT: Sleeman, Matthew  
; APPLICANT: Onrust, Rene  
; APPLICANT: Murison, James Greg  
; TITLE OF INVENTION: Compositions Isolated from Skin Cells  
; FILE REFERENCE: 11000.1011c1  
; CURRENT APPLICATION NUMBER: US/09/188,930A  
; CURRENT FILING DATE: 1998-11-09  
; NUMBER OF SEQ ID NOS: 348  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 159  
; LENGTH: 206  
; TYPE: PRT  
; ORGANISM: mouse  
; US-09-188-930-159

Query Match 29.6%; Score 346; DB 4; Length 206;  
Best Local Similarity 40.2%; Pred. No. 5.1e-29;  
Matches 80; Conservative 31; Mismatches 66; Indels 22; Gaps 8;

```

Qy 9 LVCLLVHAAFRVVGQGMQAFKNDATFIIPELGEYEP-PPLELNKKTMMRAENGRPH 67
Db 13 LVCLLMNCL-----AFKNDATEIL--YSHVVKPVPAHSSNSTLNOANGR--H 59
Qy 68 HPFETKDV---EYSCRELHFTRYVTDGPCRSAPVTELYVCSGCGPARLLPNAIGRG-- 122
Db 60 FSTGLDRNTRVOYGCRLSTKTYISDSGCTISPLKELVCAECPLPLPVLPMWIGGYG 119
Qy 123 -KWM-RPSGDPFCIPDRYRAORVOLLCPGGAAPRAKRVLVASCCKRLTRFHNOSELK 180
Db 120 TKYMSRRSSQEMRCVNDKTRTORIQLOCDGST-RTYKITVYVACKCKRYTROHNESSHN 178
Qy 181 DFGEAPARPOKGRKPRPA 199
Db 179 FESVPAKPVQHHRRKRA 197

```

RESULT 4  
US-09-188-930-286

; Sequence 286, Application US/09188930A  
; Patent No. 6150502

; GENERAL INFORMATION:

; APPLICANT: Watson, James D.  
; APPLICANT: Strachan, Lorna  
; APPLICANT: Sleeman, Matthew  
; APPLICANT: Onrust, Rene  
; APPLICANT: Murison, James Greg  
; TITLE OF INVENTION: Compositions Isolated from Skin Cells  
; FILE REFERENCE: 11000.1011c1  
; CURRENT APPLICATION NUMBER: US/09/188,930A  
; CURRENT FILING DATE: 1998-11-09  
; NUMBER OF SEQ ID NOS: 348  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 286  
; LENGTH: 206



F:9-34,23-57,26-88,38-110,72-100,90-93/Disulfide bonds: #status predicted  
F:13/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 7.9%; Score 93; DB 2; Length 118;

Best Local Similarity 30.1%; Pred. No. 0.3;

Matches 37; Conservative 13; Mismatches 55; Indels 18; Gaps 6;

QY 78 CRELHTRFLTDGPCRSAPVTELVCSGQC-GPARLLPNAIGRVKWMRPNCGDFRCIPDR 136

Db 9 CRPINATLAQZACPVCTITFTTSICAGYCPSMNVLPAL-----PPVPZPVCTYRQ 61

QY 137 YRAQRYOLL-CPGGAAPR-SRKRVLVASCKCKRLTRFHNSGLKDFGPTARPOK-GRKP 193

Db 62 LRFASIRLPCGPPGVPMVSPFVALSCHGCPRLSS-----SBGCGRAZPLACBRSP 114

QY 194 RFG 196

Db 115 RFG 117

RESULT 3

teinizizing hormone beta subunit - turkey

C:Species: Meleagris gallopavo (common turkey)

C>Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 16-Jul-1999

C:Accession: I51373

R:You, S.; Foster, L.R.; Silsby, J.L.; el Halawani, M.E.; Foster, D.N.

J. Mol. Endocrinol. 14, 117-129, 1995

A:Title: Sequence analysis of the turkey LH beta subunit and its regulation by gonadotropin

A:Reference number: I51373; MUID:95290073

A:Accession: I51373

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-159 <Y00>

A:Cross-references: GB:L35519; NID:9530952; PIDN:AAA74125.1; PID:9530953

C:Genetics:

A:Gene: LH-beta

C:Superfamily: pituitary glycoprotein hormone beta chain

Query Match

Best Local Similarity 26.6%; Pred. No. 0.45;

Matches 33; Conservative 12; Mismatches 58; Indels 21; Gaps 5;

QY 60 GGRPHHPRDAKOVSEYSCRELHYTRFLTDGPCRSAPVTELVCSGQC-GPARLLPNAIGR 119

Db 43 GGRPP-----CRPINVTAVAEKDCPCQCAVTTTACGCTCR-----EPVYR 85

QY 120 VKWMPNGDFRCIPDRYRAQRYOLL-CPGGAAPRSRKRVLVASCKCKRLTRFHNSGLK 178

Db 86 SPLGRD--PQSCTYGALRYERMALMGCPISDPRV-LLPVALSCRCARCIATSDCTVQ 142

QY 179 DRCR 182

Db 143 GUGP 146

RESULT 4

lutropin beta chain - minke whale

N:Alternate names: luteinizing hormone beta chain

C:Species: Balaeoptera acutrostrata (minke whale, lesser torqual)

C>Date: 07-May-1993 #sequence\_revision 07-May-1993 #text\_change 08-Dec-1995

C:Accession: PNO139

R:Karasev, V.S.; Pankov, Y.A.

Biochimica 50, 1972-1986, 1985

A:Title: Amino acid sequence of reduced and carboxymethylated alpha- and beta-subunits

A:Reference number: PNO138

A:Accession: PNO139

A:Molecule type: protein

A:Residues: 1-118 <KAR>

A:Note: article in Russian with English abstract

C:Superfamily: pituitary glycoprotein hormone beta chain  
C:Keywords: glycoprotein; hormone  
F:9-34,23-57,26-88,38-110,72-100,90-93/Disulfide bonds: #status predicted  
F:13/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 7.8%; Score 91; DB 2; Length 118;

Best Local Similarity 29.3%; Pred. No. 0.45;

Matches 36; Conservative 12; Mismatches 57; Indels 18; Gaps 6;

QY 78 CRELHTRFLTDGPCRSAPVTELVCSGQC-GPARLLPNAIGRVKWMRPNCGDFRCIPDR 136

Db 9 CRPINATLAQZACPVCTITFTTSICAGYCPSMNVLPAL-----PPVPZPVCTYRQ 61

QY 137 YRAQRYOLL-CPGGAAPR-SRKRVLVASCKCKRLTRFHNSGLKDFGPTARPOK-GRKP 193

Db 62 LRFASIRLPCGPPGVPMVSPFVALSCHGCPRLSS-----SBGCGRAZPLACBRSP 114

QY 194 RFG 196

Db 115 RFG 117

RESULT 5

lutropin beta chain precursor - chicken

N:Alternate names: luteinizing hormone beta chain

C:Species: Gallus gallus (chicken)

C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 08-Dec-1995

C:Accession: A61091

R:Moce, T.; Ando, H.; Ueda, T.; Kubokawa, K.; Higashinakagawa, T.; Ishii, S.

J. Mol. Endocrinol. 3, 129-137, 1989

A:Title: Molecular cloning and nucleotide sequence analysis of the putative cDNA for

A:Reference number: A61091; MUID:89374710

A:Accession: A61091

A:Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-158 <NOC>

C:Superfamily: pituitary glycoprotein hormone beta chain

C:Keywords: glycoprotein; hormone; pituitary

F:1-39/Domain: signal sequence #status predicted <SIG>

F:46-73,62-96,65-127,77-149,111-139,129-132/Disulfide bonds: #status predicted

F:52/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

Best Local Similarity 23.7%; Pred. No. 0.81;

Matches 32; Conservative 11; Mismatches 49; Indels 43; Gaps 5;

QY 60 GGRPHHPRDAKOVSEYSCRELHYTRFLTDGPCRSAPVTELVCSGQC-----G 108

Db 43 GGRPP-----CRPINVTAVAEKDCPCQCAVTTTACGCTCR-----EPVYR 89

QY 109 PARLLPNAIGRVKWMRPNCGDFRCIPDRYRAQRYOLL-CPGGAAPRSRKRVLVASCKCKR 167

Db 90 PP-----PQSCTYGALRYERMALMGCPISDPRV-LLPVALSCRCAR 131

QY 168 LTRFHNSGLKDFGR 182

Db 132 CPMAISDCTVQGLRP 146

RESULT 6

ribosomal protein L2 - evening primrose mitochondrion

C:Species: mitochondrion Oenothera villaricae (evening primrose)

C>Date: 27-Jan-1995 #sequence\_revision 24-Feb-1995 #text\_change 13-Aug-1999

C:Accession: S46947

R:Frut, I.; Schuster, W.

submitted to the EMBL data library, July 1994

A:Reference number: S46947

A:Accession: S46947

A:Molecule type: DNA

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2001, 16:01:13 ; Search time 31.13 Seconds

(without alignments)  
232.185 Million cell updates/sec

Title: US-09-668-021-12

Sequence: 1 MDPSLAPCLICLVHAAFC.....KPPRGAKAKANOAELENNY 211

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	93	7.9	118 1 LSHB_PHYCA	P25330 physeter ca
2	92.5	7.9	159 1 LSHB_MELEGA	P45646 meleagris g
3	91	7.8	118 1 LSHB_BALAC	P33088 balaenopter
4	86.5	7.4	141 1 LSHB_TRIVU	O46482 trichosurus
5	82	7.0	301 1 UL49_HSV11	P10233 herpes simp
6	81.5	7.0	2182 1 CAB1_RAT	O08480 rattus simp
7	81	6.9	470 1 KLF4_HUMAN	O43474 homo sapien
8	80.5	6.9	348 1 GCP_SYNY3	P34034 synechocyst
9	80.5	6.9	975 1 CDP_CANFA	P39881 canis fam11
10	80	6.8	141 1 LSHB_CENSI	O77833 ceratotheri
11	80	6.8	888 1 KLTR_MOUSE	P08923 mus musculu
12	79.5	6.8	166 1 LSHB_COTJA	P45657 coturnix co
13	78.5	6.7	625 1 DUS8_HUMAN	O13202 homo sapien
14	78	6.7	1942 1 Y054_HUMAN	P42694 homo sapien
15	77.5	6.6	138 1 LSHB_MACRU	O46483 macropus ru
16	77.5	6.6	188 1 VEGB_MOUSE	P49766 mus musculu
17	77.5	6.6	623 1 EXAA_PSEAE	Q92437 pseudomona
18	77	6.6	1959 1 AGR1_RAT	P25304 rattus norv
19	76.5	6.5	165 1 CGBH_HUMAN	P01233 homo sapien
20	76.5	6.5	170 1 PLGF_HUMAN	P49763 homo sapien
21	76	6.5	143 1 LSHB_FELCA	O77805 felis silve
22	76	6.5	644 1 VP4_BRV11	P33428 bluetongue
23	76	6.5	644 1 VP4_BRV13	P33428 bluetongue
24	76	6.5	1056 1 WUC5_HUMAN	P38088 homo sapien
25	74.5	6.4	128 1 LSHB_STRCA	P80664 struthio ca
26	74.5	6.4	232 1 HXB9_XENLA	P11272 xenopus lae
27	74	6.3	180 1 DAN_HUMAN	P41271 homo sapien
28	74	6.3	379 1 CYR6_MOUSE	P18406 mus musculu
29	74	6.3	381 1 CYR6_HUMAN	O00622 homo sapien
30	74	6.3	509 1 CR03_BRANA	P33523 brassica na
31	73.5	6.3	1480 1 SLIT_DROME	P24014 drosophila
32	73.5	6.3	174 1 BARI_CHITE	P02849 chironomus
33	73.5	6.3	474 1 KLF4_MOUSE	O60793 mus musculu

34	73.5	6.3	654 1 VP4_BRV10	P07132 bluetongue
35	73.5	6.3	969 1 SYL_MYCTU	P71698 mycobacteri
36	73.5	6.3	1172 1 LMB3_HUMAN	O13751 homo sapien
37	73	6.2	178 1 DAN_RAT	O06880 rattus norv
38	73	6.2	992 1 ERN6_ERV	P03204 epstein-bar
39	73	6.2	1186 1 CEN4_BACTS	Q45710 bacillus th
40	73	6.2	1955 1 AGR1_CHICK	P31696 gallus gall
41	73	6.2	2481 1 UN52_CAEL	O06561 caenorhabdi
42	72.5	6.2	628 1 MSIN_HUMAN	O13421 homo sapien
43	72	6.1	401 1 HME1_MOUSE	P09065 mus musculu
44	72	6.1	532 1 ZN20_HUMAN	P17024 homo sapien
45	72	6.1	890 1 AT8S_HUMAN	G9up79 homo sapien

## ALIGNMENTS

RESULT	1	STANDARD	PRT	118 AA
LSHB_PHYCA				
ID	LSHB_PHYCA			
AC	P25330;			
DT	01-MAY-1992 (Rel. 22, Created)			
DT	01-MAY-1992 (Rel. 22, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	LUTEOPIN BETA CHAIN (LUTEINIZING HORMONE BETA SUBUNIT) (LSH-BETA)			
DE	(LSH-B) (LH-B).			
GN	LHB.			
OS	Physeter catodon (Sperm whale) (Physeter macrocephalus).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti;			
OC	Physetridae; Physeter.			
OX	NCBI_TaxId=9755;			
RN	[1]			
RP	SEQUENCE.			
RX	MEDLINE=87032654; PubMed=3771098;			
RA	Pankov Y.A., Karasyov V.S.;			
RT	"Primary structure of sperm whale luteinizing hormone.";			
RL	Int. J. Pept. Protein Res. 28:124-129(1986).			
RM	[2]			
RP	SEQUENCE.			
RX	MEDLINE=64281133; PubMed=6466737;			
RA	Pankov Y.A., Karasev V.S.;			
RT	"Luteinizing hormone of the sperm whale. Amino acid sequences of reduced and carboxymethylated beta-subunits.";			
RL	Biochimica 49:1004-1018(1984).			
CC	-1- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.			
CC	-1- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN, LUTEOPIN, FOLLITROPIN AND GONADOTROPIN.			
CC	-1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN FAMILY.			
CC	PIR: PNO141; PNO141.			
DR	HSSP: P01233; 1HRP.			
DR	InterPro: IPR000359; -.			
DR	InterPro: IPR001545; -.			
DR	Pfam: PF00007; Cys_knot; 1.			
DR	PRINTS: PRO0438; GFCYSKNOT.			
DR	PROSITE: PS00261; GLYCO_HORMONE_BETA_1; 1.			
DR	PROSITE: PS00689; GLYCO_HORMONE_BETA_2; 1.			
KW	Hormone; Glycoprotein.			
FT	DISULFID 9	57		BY SIMILARITY.
FT	DISULFID 23	72		BY SIMILARITY.
FT	DISULFID 26	110		BY SIMILARITY.
FT	DISULFID 34	88		BY SIMILARITY.
FT	DISULFID 38	90		BY SIMILARITY.
FT	DISULFID 93	100		BY SIMILARITY.
FT	CARBOHYD 13	13		N-LINKED (GLCNAC. . .)
SQ	SEQUENCE 118 AA; 12412 MW; 8117756382F15E7 CRC64;			

Query Match 7.9%; Score 93; DB 1; Length 118;

Best Local Similarity 30.1%; Pred. No. 0.089;  
Matches 37; Conservative 13; Mismatches 55; Indels 18; Gaps 6;

QY 78 CRELHYTRFLDGPGRSAKPYTELVCSGOC-GPARLLPNAIGRVMKRNPGDFRCIDPR 136  
DB 9 CRINMTLANQACACVCTFTTSTICAGCPSMVRVLPAL-----PPVZPCTYRQ 61  
QY 137 YRAORVOLL-CPGGAAPR-SRKVRLVASCCKRLTRFNHOSLKDGPETAPRQK-GRKP 193  
DB 62 LRFASIRLPGCPGVNPNVSPFVALSCHGCPRLS-----SDCGPGRADPLACNRSP 114  
QY 194 RFG 196  
DB 115 RFG 117

## RESULT 2

LSHB\_MELGA  
ID LSHB\_MELGA STANDARD; PRT; 159 AA.  
P45646;

01-NOV-1995 (Rel. 32, Created)  
01-NOV-1995 (Rel. 32, Last sequence update)  
01-OCT-2000 (Rel. 40, Last annotation update)  
LUTROPIN BETA CHAIN PRECURSOR (LUTEINIZING HORMONE BETA SUBUNIT) (LSH-BETA) (LSH-B) (LH-B).

OS Melagris gallopavo (Common turkey).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Meleagridae; Meleagris.  
OX NCBI\_Taxid=9103;  
RN [1]

RP SEQUENCE FROM N.A.  
RC TISSUE=Plutary;  
RX MEDLINE=95290073; PubMed=7772235;  
RA You S., Foster L.K., Sillsby J.L., el Halawani M.E., Foster D.N.;  
RT Sequence analysis of the turkey LH beta subunit and its regulation  
RT by gonadotrophin-releasing hormone and prolactin in cultured  
RT pituitary cells.";

RL J. Mol. Endocrinol. 14:117-129(1995).

-1- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING  
THE TESTES AND OVARIES TO SYNTHESIZE SPEROIDS.

-1- SUBUNIT: HETEROIDIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA  
CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,  
LUTROPIN, FOLLITROPIN AND GONADOTROPIN.

-1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN  
FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: L35519; AAA74125.1; ALT\_INT.

DR HSSP: P01233; IHRP.

DR InterPro: IPR000359; -

DR InterPro: IPR001545; -

DR Pfam: PF00007; Cys\_knot; 1.

DR PRINTS: PR00438; GFCYSKNOT.

DR PROSITE: PS00261; GLYCO\_HORMONE\_BETA\_1; 1.

DR PROSITE: PS00689; GLYCO\_HORMONE\_BETA\_2; 1.

KW Hormone; Signal; Glycoprotein.

FT SIGNAL 1 39  
FT CHAIN 40 159  
FT DISULFID 48 96  
FT DISULFID 62 111  
FT DISULFID 65 149  
FT DISULFID 73 127  
FT DISULFID 77 129  
FT DISULFID 132 139

FT CARBOHYD 52 52 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 159 AA; 16285 MW; 52B50C8C879653C6 CRC64;

Query Match 7.9%; Score 92.5; DB 1; Length 159;

Best Local Similarity 26.6%; Pred. No. 0.13;  
Matches 33; Conservative 12; Mismatches 58; Indels 21; Gaps 5;

QY 60 GGRPHHHPYDAKVSEYSCRELHYTRFLDGPGRSAKPYTELVCSGOC-GPARLLPNAIGR 119  
DB 43 GGRPP-----CRPINVTVAVEKDECPQCMATVTAACGGYCRTR----EPYR 85

QY 120 VKMWRPNPGRFCIPDRYRAORVOLL-CPGGAAPRSRKRLVASCCKRLTRFNHOSLKD 178  
DB 86 SPLGRP--POSSCTYGALRYERWALMGCPIGSDPRV-LLPVALSCRCARCPATSDCTVQ 142

QY 179 DFGP 182  
DB 143 GLGP 146

## RESULT 3

LSHB\_BALAC  
ID LSHB\_BALAC STANDARD; PRT; 118 AA.  
P33088;

01-OCT-1993 (Rel. 27, Created)

01-OCT-1993 (Rel. 27, Last sequence update)

01-OCT-2000 (Rel. 40, Last annotation update)  
LUTROPIN BETA CHAIN (LUTEINIZING HORMONE BETA SUBUNIT) (LSH-BETA)  
(LSH-B) (LH-B).

OS Balaeoptera acutirostrata (Mink whale) (lesser rorqual).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Mysticeti;  
OC Balaeopteridae; Balaeoptera.  
OX NCBI\_Taxid=9767;  
RN [1]

RP SEQUENCE.

RA Karasev V.S., Pankov Y.A.;  
RT "amino acid sequence of reduced and carboxymethylated alpha- and beta-  
RT subunits of the little picked whale luteinizing hormone.";

RL Biochimica 50:1972-1986(1985).

-1- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING  
THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.

-1- SUBUNIT: HETEROIDIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA  
CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,  
LUTROPIN, FOLLITROPIN AND GONADOTROPIN.

-1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN  
FAMILY.

DR PIR: P0135; P0139.

DR HSSP: P01233; IHRP.

DR InterPro: IPR000359; -

DR InterPro: IPR001545; -

DR Pfam: PF00007; Cys\_knot; 1.

DR PRINTS: PR00438; GFCYSKNOT.

DR PROSITE: PS00261; GLYCO\_HORMONE\_BETA\_1; 1.

DR PROSITE: PS00689; GLYCO\_HORMONE\_BETA\_2; FALSE\_NEG.

KW Hormone; Glycoprotein.

FT DISULFID 9 57  
FT DISULFID 23 72  
FT DISULFID 26 110  
FT DISULFID 34 88  
FT DISULFID 38 90  
FT DISULFID 93 100  
FT CARBOHYD 13 13

FT SEQUENCE 118 AA; 12414 MW; 039F229EFC480F5D CRC64;

Query Match 7.8%; Score 91; DB 1; Length 118;  
Best Local Similarity 29.3%; Pred. No. 0.13;  
Matches 36; Conservative 12; Mismatches 57; Indels 18; Gaps 6;



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2001, 16:00:35 ; Search time 89.34 Seconds  
(without alignments)  
312.473 Million cell updates/sec

Title: US-09-668-021-12

Perfect score: 1171  
Sequence: 1 MQPSLAPCICLCILYHAFFCA.....KRPFGARGAKANQAELENNY 211

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

ached: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_16:\*  
1: sp.archaea:\*  
2: sp.bacteria:\*  
3: sp.fungi:\*  
4: sp.human:\*  
5: sp.invertebrate:\*  
6: sp.mammal:\*  
7: sp.mhc:\*  
8: sp.organelle:\*  
9: sp.phage:\*  
10: sp.plant:\*  
11: sp rodent:\*  
12: sp\_unclassified:\*  
13: sp.vertebrate:\*  
14: sp\_virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	249.5	21.3	134	4 Q9Y303	Q9Y303 homo sapien
2	90.5	7.7	184	3 073755	Q73755 gallus gall
3	90	7.7	712	4 Q9Y557	Q9Y557 homo sapien
4	90	7.7	720	4 Q9NY23	Q9NY23 homo sapien
5	90	7.7	988	6 Q97867	Q97867 sus scrofa
6	89	7.6	215	8 Q9MFB6	Q9MFB6 beta vulgar
7	89	7.6	261	2 Q9F1E4	Q9F1E4 acetobacter
8	89	7.6	262	13 013091	013091 pleurodeles
9	88.5	7.6	182	13 073754	073754 xenopus lae
10	88.5	7.6	184	4 060565	060565 homo sapien
11	88	7.5	184	11 035793	035793 ratus norv
12	88	7.5	281	2 Q9EMD9	Q9EMD9 streptomyc
13	88	7.5	332	10 Q40600	Q40600 oenothera b
14	87	7.4	272	13 Q9PWB0	Q9PWB0 gallus gall
15	86.5	7.4	482	14 Q71654	Q71654 human immun
16	86.5	7.4	1114	11 Q9JKW7	Q9JKW7 mus musculu
17	86	7.3	184	11 070326	070326 mus musculu
18	85	7.3	168	4 Q9H772	Q9H772 homo sapien
19	85	7.3	272	13 Q9PUK2	Q9PUK2 gallus gall

20	84.5	7.2	791	11 070368	070368 ratus norv
21	84.5	7.2	1223	11 090W33	090W33 ratus sp.
22	84	7.2	168	11 088273	088273 mus musculu
23	83	7.1	1035	5 Q9NEG1	Q9NEG1 drosophila
24	82.5	7.0	334	10 049658	049658 arabidopsis
25	82	7.0	1081	4 076065	076065 homo sapien
26	82	7.0	1537	11 Q9TW6	Q9TW6 mus musculu
27	82	7.0	1581	11 Q9R0B6	Q9R0B6 mus musculu
28	82	7.0	1637	6 Q9XSV8	Q9XSV8 bos taurus
29	81.5	7.0	349	10 P93311	P93311 arabidopsis
30	81.5	7.0	13288	6 018758	018758 sus scrofa
31	81	6.9	470	4 Q9UNP3	Q9UNP3 homo sapien
32	81	6.9	1242	4 Q9NS15	Q9NS15 homo sapien
33	81	6.9	1382	4 Q9H7K2	Q9H7K2 homo sapien
34	80.5	6.9	270	13 P70041	P70041 xenopus lae
35	80.5	6.9	1704	5 Q94446	Q94446 chironomus
36	80	6.8	267	4 Q95813	Q95813 homo sapien
37	80	6.8	866	5 P91426	P91426 caenorhabdi
38	80	6.8	1212	2 Q9LIC8	Q9LIC8 streptomyc
39	80	6.8	3161	2 054511	054511 yerinia en
40	79.5	6.7	1616	4 015054	015054 homo sapien
41	78.5	6.7	641	5 Q97355	Q97355 drosophila
42	78.5	6.7	641	5 Q9YVW6	Q9YVW6 drosophila
43	78.5	6.7	648	4 Q9UP38	Q9UP38 homo sapien
44	78.5	6.7	818	4 Q9Z0M6	Q9Z0M6 mus musculu
45	78.5	6.7	1142	5 Q26615	Q26615 strongyloce

#### ALIGNMENTS

RESULT 1					
Q9Y303	Q9Y303	PRELIMINARY:	PRT:	134 AA.	
AC	Q9Y303;				
DT	01-NOV-1999 (TREMBLrel. 12, Created)				
DT	01-NOV-1999 (TREMBLrel. 12, Last sequence update)				
DT	01-MAY-2000 (TREMBLrel. 13, Last annotation update)				
DE	HYPOTHEITICAL 15.3 KDA PROTEIN (FRAGMENT).				
GN	DKZP564D206.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=BRAIN;				
RA	Wambutt R., Heubner D., Mewes H.W., Gassenhuber J., Wiemann S.;				
RL	Submitted (MAY-1999) to the EMBL/Genbank/DBJ databases.				
DR	EMBL; AL050024; CAB43243.1; -.				
DR	InterPro; IPR000359; -.				
DR	PROSITE; PS01225; CTCK_2; 1.				
KW	Hypothetical protein.				
FT	NON_TER				
SO	SEQUENCE	134 AA;	15324 MW;	A0B7A8650DZEE6F1	CRC64;

Query Match 21.3%; Score 249.5; DB 4; Length 134;  
Best Local Similarity 39.7%; Pred. No. 2.3e-17;  
Matches 52; Conservative 28; Mismatches 44; Indels 7; Gaps 4;

QY	78 CRELIHYRFLTDGPCRSKRPVTELVCSGCGPARALPNAIG---RYKMW-RPAGPDPFRCI 133	
DB	3 CRELIHSTFYISDGOCTSIPLKELVCAEGECILLPVLNWTGGGYKRWSSQSWRCV 62	
QY	134 PDRRAORVOLLCGGAAPRSRKVRLVASCKKRLTFPHNOSLELKDGPRTAPRQGRKP 193	
DB	63 NDKRTTGQIQLQCODGSR-RYKKTIVYTACKCKRYTQHNHSSINFPSSMSPARFVQHNRE 121	
QY	194 RFGARGAANQ 204	
DB	122 R--KRASKSK 130	



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2001, 15:48:42 ; Search time 80.28 Seconds  
(without alignments)  
159.338 Million cell updates/sec

Title: US-09-668-021-12

Perfect score: 1171  
Sequence: 1 MGPLAPLCICLLVHAHFCA.....KPRPGARKANQALENAVY 211

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Indexed: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_0601.\*  
1: /SID58/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID58/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID58/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID58/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID58/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID58/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID58/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID58/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID58/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID58/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID58/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID58/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID58/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID58/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID58/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID58/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID58/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID58/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID58/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID58/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID58/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID58/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1171	100.0	211	21	AAV96432 Murine TGF-beta b1
2	1094	93.4	213	21	AAV96433 Rat TGF-beta bindi
3	1052	89.8	213	21	AAV96431 Vervet TGF-beta b1
4	1029	87.9	213	21	AAV96430 Human DAN/Cerberus
5	1029	87.9	213	21	AAV96429 Human TGF-beta bin
6	1029	87.9	213	22	AAV97589 Human secreted pro
7	1021	87.2	213	21	AAV96430 Human TGF-beta bin
8	1020	87.1	213	21	AAV96436 Human TGF-beta bin
9	934	79.8	176	21	AAV96434 Bovine DAN/Cerberus
10	857.5	73.2	206	21	AAV96434 Bovine TGF-beta b1
11	340.5	29.1	206	21	AAV95981 Murine skin cell p

12	340.5	29.1	206	21	AAV76031
13	340.5	29.1	206	22	AAV55920
14	340.5	29.1	206	22	AAV55970
15	333	28.4	206	18	AAV05408
16	333	28.4	206	19	AAV58704
17	333	28.4	206	20	AAV95711
18	333	28.4	206	21	AAV10233
19	215	18.4	50	20	AAV12009
20	114	9.7	116	18	AAV27654
21	114	9.7	116	18	AAV44090
22	98.5	8.4	102	20	AAV03225
23	98.5	8.4	102	21	AAV10277
24	90.5	7.7	184	19	AAV29727
25	90	7.7	712	21	AAV43147
26	88.5	7.6	182	19	AAV29726
27	88.5	7.6	184	19	AAV6293
28	88.5	7.6	184	20	AAV42173
29	88.5	7.6	184	21	AAV95961
30	88	7.5	184	20	AAV42174
31	88	7.5	429	20	AAV42175
32	87	7.4	391	20	AAV42178
33	86	7.3	184	20	AAV42172
34	86	7.3	212	21	AAV43169
35	86	7.3	301	21	AAV83261
36	85	7.3	168	21	AAV84014
37	84	7.2	168	21	AAV51132
38	83.5	7.1	165	20	AAV99515
39	83	7.1	386	18	AAV18664
40	82.5	7.0	145	12	AAV15109
41	82.5	7.0	304	21	AAV20509
42	82.5	7.0	304	21	AAV35061
43	82.5	7.0	334	21	AAV20508
44	82.5	7.0	334	21	AAV35060
45	82	7.0	258	22	AAV88367

#### ALIGNMENTS

RESULT	1
ID	AAV96432 standard; Protein; 211 AA.
XX	
AC	AAV96432;
DT	12-SEP-2000 (first entry)
XX	
DE	Murine TGF-beta binding protein (BEER).
KW	osteopathic; transforming growth factor-beta; TGF-beta; binding protein; BEER; gene therapy; antisense therapy; fracture; bone mineralization.
XX	
OS	Mus musculus.
XX	
PN	WO200032773-A1.
XX	
PD	08-JUN-2000.
PF	24-NOV-1999; 99WO-US27990.
XX	
PR	27-NOV-1998; 98US-0110283.
XX	
PA	(DARW-) DARWIN DISCOVERY LTD.
PI	Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepker BW;
PI	Van Ness J, Winkler DG;
XX	
DR	WPI; 2000-412321/35.
XX	
DR	N-PSDB; AAA29058.
PT	Nucleic acids (I) encoding a transforming growth factor beta binding protein, useful for identifying agents for treating osteopenia,
PT	osteoporosis and fractures

Murine skin cell p  
Skin cell protein,  
Skin cell protein,  
Human small CCN-11  
Human small CCN-11  
Homo sapiens fetal  
Human adult retina  
Human 5' EST secre  
Secreted protein A  
Human secreted pro  
Amino acid sequenc  
Human fetal kidney  
DAN and b57 protei  
Human ORFX ORF2911  
DAN and b57 protei  
Human b57 protei  
Human b57 protein  
Mouse DRM protein  
Human gremlin. Ho  
Rat DRM protein se  
EGFP/DRM fusion pr  
EGFP/DRM fusion pr  
Human DRM protein  
Human ORFX ORF2933  
HSV-1 V22 cellular  
Amino acid sequenc  
Murine cerebral ne  
Glycoprotein hormo  
Fragmented human N  
hCG/ELH chimera, E  
Arabidopsis thalia  
Arabidopsis thalia  
Arabidopsis thalia  
Arabidopsis thalia  
Human membrane or

XX Claim 5; Page 124; 162pp; English.  
 PS  
 CC This shows a murine transforming growth factor-beta (TGF-beta)  
 CC binding protein designated mBER. The cDNA and protein may be used for  
 CC prevention, treatment and diagnosis of diseases associated with  
 CC inappropriate BEER expression. For example, they may be used to treat  
 CC disorders associated with decreased TGF-beta BP expression. The cDNA or  
 CC vectors may be administered to treat diseases by rectifying mutations or  
 CC deletions in a patient's genome that affect the activity of BEER by  
 CC expressing inactive proteins or to supplement the patients own production  
 CC of BEER polypeptides. The nucleic acids may be used for recombinant  
 CC production of BEER, gene therapy, antisense therapy, as probes for  
 CC diagnostic assays and for functional studies. BEER may be used to raise  
 CC antibodies and for identification of BEER modulators. BEER antagonists  
 CC may be used to increase bone mineral content for the treatment of  
 CC disorders such as osteopenia, osteoporosis, fractures and other disorders  
 CC associated with low mineral content.  
 CC  
 SQ Sequence 211 AA;

Query Match 100.0%; Score 1171; DB 21; Length 211;  
 Best Local Similarity 100.0%; Pred. No. 5.5e-100;  
 Matches 211; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOPS LAPCLICILVHAFC AVEGOGQAFRNDATEV I PGLGEYPPPPNNQTMNRAENG 60  
 Db 1 mgsplapclilcllvha favegvgwqafndatevlp9lgeyppppennqumraeng 60  
 QY 61 GRP HHPRYPAKDVSEYSCRELHTRFLTDGPCRSAPRYVELVCSGCGPARLLPNAIGRV 120  
 Db 61 grpphpyakdvseyscrelhytrfltdgpcrsakprvelvcsqcgparllpnaigrv 120  
 QY 121 KKKMRPNPDPFRCLPDRYRAQROVOLLCPGGAAPRSKRVRLVASCCKRLTRFNHOSLEKPF 180  
 Db 121 kkkmrpnppdfrclpdr yraqrvql lcp9gaaprskrvrlvasckckrltrfnhgselkpf 180  
 QY 181 GPETARPQGRKRPARGAKAQAQAELENAY 211  
 Db 181 gpetar pqkyrkprp argakanaqaelenay 211

## RESULT 2

AA96433  
 ID AAY96433 standard; Protein: 213 AA.

AC AAY96433;

DT 12-SEP-2000 (first entry)

XX Rat TGF-beta binding protein (BEER).

KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;  
 KM BEER; gene therapy; antisense therapy; fracture; bone mineralization.

OS Rattus norvegicus.

PN MO200032773-A1.

XX 08-JUN-2000.

PD 24-NOV-1999; 99WO-US27990.

PF 27-NOV-1998; 98US-0110283.

XX (DARW-) DARWIN DISCOVERY LTD.

PA Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;

PI Van Ness J, Winkler DG;

XX WPI; 2000-412321/35.

DR N-PSDB; AAA29059.

XX Nucleic acids (1) encoding a transforming growth factor beta binding  
 PT protein, useful for identifying agents for treating osteopenia,  
 PT osteoporosis and fractures  
 PS  
 CC Claim 6; Page 125-126; 162pp; English.

XX This shows a rat transforming growth factor-beta (TGF-beta) binding  
 CC protein designated rBER. The cDNA and protein may be used for  
 CC prevention, treatment and diagnosis of diseases associated with  
 CC inappropriate BEER expression. For example, they may be used to treat  
 CC disorders associated with decreased TGF-beta BP expression. The cDNA or  
 CC vectors may be administered to treat diseases by rectifying mutations or  
 CC deletions in a patient's genome that affect the activity of BEER by  
 CC expressing inactive proteins or to supplement the patients own production  
 CC of BEER polypeptides. The nucleic acids may be used for recombinant  
 CC production of BEER, gene therapy, antisense therapy, as probes for  
 CC diagnostic assays and for functional studies. BEER may be used to raise  
 CC antibodies and for identification of BEER modulators. BEER antagonists  
 CC may be used to increase bone mineral content for the treatment of  
 CC disorders such as osteopenia, osteoporosis, fractures and other disorders  
 CC associated with low mineral content.  
 CC  
 SQ Sequence 213 AA;

Query Match 93.4%; Score 1094; DB 21; Length 213;  
 Best Local Similarity 93.9%; Pred. No. 6.5e-93;  
 Matches 200; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

QY 1 MOPS LAPCLICILVHAFC AVEGOGQAFRNDATEV I PGLGEYPPPP--ENNQTMNRAE 58  
 Db 1 mgsplapclilcllvha favegvgwqafndatevlp9lgeyppppqelnqumrae 60  
 QY 59 NGRP HHPRYPAKDVSEYSCRELHTRFLTDGPCRSAPRYVELVCSGCGPARLLPNAIG 118  
 Db 61 ngrpphpyakdvseyscrelhytrfltdgpcrsakprvelvcsqcgparllpnaig 120  
 QY 119 RKKMRPNPDPFRCLPDRYRAQROVOLLCPGGAAPRSKRVRLVASCCKRLTRFNHOSLEK 178  
 Db 121 rkkmrpnppdfrclpdr yraqrvql lcp9gaaprskrvrlvasckckrltrfnhgselk 180  
 QY 179 DFGPETARPQGRKRPARGAKAQAQAELENAY 211  
 Db 181 dfgpetar pqkyrkprp argakanaqaelenay 213

## RESULT 3

AA96431  
 ID AAY96431 standard; Protein: 213 AA.

AC AAY96431;

DT 12-SEP-2000 (first entry)

XX Vervet TGF-beta binding protein (BEER).

KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;  
 KM BEER; gene therapy; antisense therapy; fracture; bone mineralization.

OS Cercopithecus pygerythrus.

PN MO200032773-A1.

XX 08-JUN-2000.

PD 24-NOV-1999; 99WO-US27990.

PF 27-NOV-1998; 98US-0110283.

XX (DARW-) DARWIN DISCOVERY LTD.

PA Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;



APPLICANT: Watson, James D.  
APPLICANT: Strachan, Lorna  
APPLICANT: Sleeman, Matthew  
APPLICANT: Orrust, Rene  
APPLICANT: Morrison, James Greg  
TITLE OF INVENTION: Compositions Isolated From Skin Cells  
FILE REFERENCE: 11000.1011cl  
CURRENT APPLICATION NUMBER: US/09/188.930A  
NUMBER OF SEQ ID NOS: 348  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 286  
LENGTH: 206  
TYPE: PRP  
ORGANISM: Mouse  
US-09-188-930-286

## Query Match

29.1%; Score 340.5; DB 4; Length 206;

Best Local Similarity 37.7%; Pred. No. 8.9e-28; Mismatches 71; Indels 23; Gaps 9;

Matches 80; Conservative 38; Mismatches 71; Indels 23; Gaps 9;

4 SLAPCLICLVHAACVAGCGQAFRNDATFVIGLGEYPEPP-PENNOTMRAENGGR 62

9 SLIP-LCLIMRNC-----AFKNDATFELVSHVVKVPAHPSSNSTLQARNGR 58

63 PRHP---YDADVSEYSCRELHYRFLTDGPCNSAKPVTELVCSCGCGPARLLPNAIG- 118

59 --HFSTGLDRMSRVQVCGRELSTKYISDQCTISPLKELVCAGCECLPVPWVGIG 116

119 --RVKMW-PPNGDFRCIPDRYRAQVOLLCPGGAAPRSKRVRLVASCCKRLTRFNOS 175

117 GYGTKWTMSRSSQEMRCVMDKTRTORIOLQODGST-RTYITVYVYACKCKRTTQHNES 175

176 ELKDPETAPRQKGRKKPPGAKAKANOAE 207

176 SHNESVSPAKPQOHRER--KRASKSKHSL 205

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 206 AMINO ACIDS  
TYPE: AMINO ACID  
STRANDEDNESS:  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PROTEIN  
US-08-468-847B-2

## Query Match

28.4%; Score 333; DB 1; Length 206;

Best Local Similarity 37.3%; Pred. No. 5.3e-27; Mismatches 76; Conservative 39; Mismatches 67; Indels 22; Gaps 8;

Matches 76; Conservative 39; Mismatches 67; Indels 22; Gaps 8;

9 LICLVHAACVAGCGQAFRNDATFVIGLGEYPEPP-PENNOTMRAENGGRPHHP 67

13 LACILMKSCL-----AFKNDATFELVSHVVKVPAHPSSNSTLQARNGR--HFS 61

68 YDADVSEYSCRELHYRFLTDGPCNSAKPVTELVCSCGCGPARLLPNAIG---RVK 121

62 NTGLDRMTRVQVCGRELSTKYISDQCTISPLKELVCAGCECLPVPWVGIGYGT 121

122 MW-PPNGDFRCIPDRYRAQVOLLCPGGAAPRSKRVRLVASCCKRLTRFNOS 180

122 YMSRSSQEMRCVMDKTRTORIOLQODGST-RTYITVYVYACKCKRTTQHNES 180

181 GPETAPRQKGRKKPPGAKAKANO 204

181 SMSPAKPVQOHRER--KRASKSK 202

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2001, 15:49:47 ; Search time 54.26 Seconds  
(without alignments)  
299.027 Million cell updates/sec

Title: US-09-668-021-14

Perfect score: 1167  
Sequence: 1 MQLSLAPCLACLLVNAFAVA.....KRPRAKCAKNAQALENNAY.213

Scoring table: BLOSUM62  
Gapex 10.0 , Gapext 0.5

rched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	256.5	22.0	134	2 T08710	hypothetical prote
2	95.5	8.2	159	2 I51373	luteinizing hormo
3	92.5	7.9	158	2 A61091	lutropin beta cha
4	88	7.5	118	2 PNO141	lutropin beta cha
5	86	7.4	118	2 PNO139	lutropin beta cha
6	82.5	7.1	166	2 I51242	luteinizing hormo
7	82.5	7.1	1469	2 B36665	silt protein 2 pre
8	81	6.9	3163	2 T17440	probable polyketid
9	80	6.9	601	2 T22025	hypothetical prote
10	80	6.9	2142	2 B35098	MHC class IIT hist
11	80	6.9	3161	2 T30342	protein HMP1 - ye
12	79.5	6.8	165	1 KTH0B	choriogonadotropin
13	78.5	6.7	2098	2 T18397	protein CTRP - mal
14	78	6.7	343	2 G35070	apolipoprotein H-r
15	77.5	6.6	188	2 J04680	vascular endotheli
16	77.5	6.6	270	2 S71793	head-inducing fact
17	77.5	6.6	365	2 S67507	morphogen lefty pr
18	77	6.6	385	2 A54039	major surface prot
19	77	6.6	1480	2 A36665	silt protein 1 pre
20	77	6.6	1682	1 C70588	probable mbte prot
21	76.5	6.6	145	2 I37231	beta-gonadotropin
22	76.5	6.6	301	1 RMBEF9	U49 protein - hum
23	76.5	6.6	1223	2 S29717	adenylate cyclase
24	76	6.5	269	2 D75631	iron ABC transport
25	76	6.5	765	2 T49592	neurofilament trip
26	76	6.5	1589	2 T42233	submaxillary mucin
27	75.5	6.5	128	2 S74085	lutropin beta cha
28	75.5	6.5	705	2 T22201	hypothetical prote
29	75	6.4	379	2 A35669	gene CYR61 protein

30	75	6.4	383	2 A23516	Balbiani ring 1 ch
31	75	6.4	405	2 S33601	cell adhesion mole
32	75	6.4	450	2 G01158	tyrosine kinase ac
33	74.5	6.4	322	1 I58195	y box-binding prot
34	74.5	6.4	866	2 T29197	hypothetical prote
35	74	6.3	174	2 B0IC1T	Balbiani ring-1 ch
36	74	6.3	563	2 A36054	mucin homolog - bo
37	74	6.3	721	2 T45495	probable transposa
38	73.5	6.3	270	2 A83501	conserved hypochet
39	73.5	6.3	324	1 J39382	y box-binding prot
40	73.5	6.3	324	1 J02292	y box-binding prot
41	73.5	6.3	1042	2 A57534	mucin 5AC (clone L
42	73.5	6.3	1436	2 A46496	antigen WC1.1 prec
43	73.5	6.3	3133	2 S52093	hemocytin - silkw
44	73	6.3	305	1 S22313	B box-binding prot
45	73	6.3	332	2 S46947	ribosomal protein

## ALIGNMENTS

## RESULT 1

T08710  
hypothetical protein DKFZp564D206.1 - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 13-Aug-1999

C:Accession: T08710

R:Mambull, R.; Heudner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, March 1999

A:Reference number: Z16471

A:Accession: T08710

A:Molecule type: mRNA

A:Residues: 1-134 <WAM>

A:Cross-references: EMBL:AL050024

A:Experimental source: fetal brain; clone DKFZp564D206

C:Genetics:

A:Note: DKFZp564D206.1

Query Match 22.0% Score 256.5; DB 2; Length 134;

Best Local Similarity 41.9%; Pred. No. 8.1e-16;

Matches 52; Conservative 24; Mismatches 43; Indels 5; Gaps 3;

QY 80 CRELHYTRVTDGPCRSAPVTELVCSGCGPARLLPNAIG---RVKWW-RPNQDFRCI 135

Db 3 CRELSTRKYSIDGQCTSIPLKELVACGECILLPVLPMWIGGCGTKYWSRSSQEMRCV 62

QY 136 PDXYRAQRYQLCPGGAARSRKRVLYASCKKRLTRPHNOSELDPOPEYARPOKGRKP 195

Db 63 NDKFTQRRLQLOQDQST-RYKRTIVYACKCKRYTROHNESSHNFEESMPAKPVQHHRE 121

QY 196 RPRRA 199

Db 122 RKRA 125

RESULT 2

I51373

luteinizing hormone beta subunit - turkey

C:Species: Melalegris gallopavo (Common turkey)

C>Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 16-Jul-1999

C:Accession: I51373

R:You, S.; Foster, L.K.; Silsby, J.L.; el Halawani, M.E.; Foster, D.N.

J. Mol. Endocrinol. 14, 117-129, 1995

A:Title: Sequence analysis of the turkey LH beta subunit and its regulation by gonado

A:Reference number: I51373; MUID:95290073

A:Accession: I51373

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-159 <YOD>

A:Cross-references: GB:I35519; NID:9530952; PIDN:AAA74125.1; PID:9530953

C:Genetics:

A:Gene: LH-beta

C:Superfamily: pituitary glycoprotein hormone beta chain

Query Match 8.2%; Score 95.5; DB 2; Length 159;  
Best Local Similarity 27.4%; Pred. No. 0.19;  
Matches 34; Conservative 11; Mismatches 58; Indels 21; Gaps 5;

QY 62 GGRPHHPYDTKDVSEYSCRELHYTRFYVDGRCRSKAKPYTELVCSGQC-----G 110  
|||  
DB 43 GGRHP-----CRPINVTVAVEKDCPCQMAVTTTACGCGYCTR---EPYR 85  
QY 122 VKMWRNGDFRCIPDRYRAQRVOLL-CPGGAAPRSKRVLVASCCKRLTRFHNSQLK 180  
|||  
DB 86 SPLGRF--FQSSCTGALRYERLWALMGCPIGSDPRV-LLPVALSRCRCARPLATSDCIYQ 142  
QY 181 DEGP 184  
||  
DB 143 GLGP 146

RESULT 3

lutropin beta chain precursor - chicken

N:Alternate names: luteinizing hormone beta chain  
C:Species: Gallus gallus (chicken)  
C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 08-Dec-1995  
C:Accession: A61091

R:Note: T.; Ando, H.; Ueda, T.; Kubokawa, K.; Hiyashinakagawa, T.; Ishii, S.  
J. Mol. Endocrinol. 3, 129-137, 1989  
A:Title: Molecular cloning and nucleotide sequence analysis of the putative cDNA for the  
A:Reference number: A61091; MUID:89374710  
A:Accession: A61091

A:Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-158 <NOC>

C:Superfamily: pituitary glycoprotein hormone beta chain

C:Keywords: glycoprotein; hormone; pituitary  
F:1-39/Domain: signal sequence #status predicted <SIG>  
F:48-73,62-96,65-127,77-149,111-139,129-132/Disulfide bonds: #status predicted  
F:32/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 7.9%; Score 92.5; DB 2; Length 158;  
Best Local Similarity 24.4%; Pred. No. 0.36;  
Matches 33; Conservative 10; Mismatches 49; Indels 43; Gaps 5;

QY 62 GGRPHHPYDTKDVSEYSCRELHYTRFYVDGRCRSKAKPYTELVCSGQC-----G 110  
|||  
DB 43 GGRHP-----CRPINVTVAVEKDCPCQMAVTTTACGCGYCTRREPVRSPUG 89  
QY 111 PARLLPNAIGRVKWRBNGPDRFCIPDRYRAQRVOLL-CPGGAAPRSKRVLVASCCKR 169  
|||  
DB 90 PR-----FQSACTYGALRYERLWALMGCPIGSDPRV-LLPVALSRCRCAR 131  
QY 170 LTRFHNSQLKDEGP 184  
||  
DB 132 CPNATSDCTVOGLGP 146

RESULT 4

PN0141

lutropin beta chain - sperm whale

N:Alternate names: luteinizing hormone beta chain

C:Species: Physeter catodon (sperm whale)  
C>Date: 07-May-1993 #sequence\_revision 07-May-1993 #text\_change 07-May-1999  
C:Accession: PN0141

R:Pankov, Y.A.; Karasev, V.S.  
Biochimica 49, 1004-1018, 1984  
A:Title: Luteinizing hormone of the sperm whale: amino acid sequence of reduced and car

A:Reference number: PN0141; MUID:84281133  
A:Accession: PN0141  
A:Molecule type: protein  
A:Residues: 1-118 <PAN>

A:Note: article in Russian with English abstract

C:Superfamily: pituitary glycoprotein hormone beta chain

C:Keywords: glycoprotein; hormone  
F:9-34,23-57,26-88,38-110,72-100,90-93/Disulfide bonds: #status predicted  
F:13/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 7.5%; Score 88; DB 2; Length 118;  
Best Local Similarity 29.5%; Pred. No. 0.67;  
Matches 36; Conservative 13; Mismatches 55; Indels 18; Gaps 6;

QY 80 CRELHYTRFYVDGRCRSKAKPYTELVCSGQC-GPARLLPNAIGRVKWRBNGPDRFCIPDR 138  
|||  
DB 9 CRPINATLAQNZACPCVCIFFTTISICAGYCPSMVRVLPAL-----PVPZPVCTYRQ 61  
QY 139 YRAQRVOLL-CPGGAAPR-SRKVRVLVASCCKRLTRFHNSQLKDEGPETARPOK-GRKP 195  
|||  
DB 62 LRFASIRLPQCPGVNPNVSPFVALSCHGCPRLSS-----SDGPGRAOPLACNRSP 114  
QY 196 RP 197  
||  
DB 115 RP 116

RESULT 5

lutropin beta chain - minke whale

N:Alternate names: luteinizing hormone beta chain

C:Species: Balaenoptera acutirostrata (minke whale, lesser rorqual)  
C>Date: 07-May-1993 #sequence\_revision 07-May-1993 #text\_change 08-Dec-1995  
C:Accession: PN0139

R:Karasev, V.S.; Pankov, Y.A.  
Biochimica 50, 1972-1986, 1985  
A:Title: Amino acid sequence of reduced and carboxymethylated alpha- and beta-subunit

A:Reference number: PN0138  
A:Accession: PN0139  
A:Molecule type: protein

A:Residues: 1-118 <KAR>

A:Note: article in Russian with English abstract

C:Superfamily: pituitary glycoprotein hormone beta chain

C:Keywords: glycoprotein; hormone  
F:9-34,23-57,26-88,38-110,72-100,90-93/Disulfide bonds: #status predicted  
F:13/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 7.4%; Score 86; DB 2; Length 118;  
Best Local Similarity 28.7%; Pred. No. 1;  
Matches 35; Conservative 12; Mismatches 57; Indels 18; Gaps 6;

QY 80 CRELHYTRFYVDGRCRSKAKPYTELVCSGQC-GPARLLPNAIGRVKWRBNGPDRFCIPDR 138  
|||  
DB 9 CRPINATLAQNZACPCVCIFFTTISICAGYCPSMVRVLPAL-----PVPZPVCTYRQ 61  
QY 139 YRAQRVOLL-CPGGAAPR-SRKVRVLVASCCKRLTRFHNSQLKDEGPETARPOK-GRKP 195  
|||  
DB 62 LRFASIRLPQCPGVNPNVSPFVALSCHGCPRLSS-----SBCGPGRAZPLACBRSP 114  
QY 196 RP 197  
||  
DB 115 RP 116

RESULT 6

151242

luteinizing hormone beta-subunit - quail

C:Species: Coturnix coturnix (quail)  
C>Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 16-Jul-1999  
C:Accession: 151242

R:Ando, H.; Ishii, S.  
Gen. Comp. Endocrinol. 93, 357-368, 1994

A:Title: Molecular cloning of complementary deoxyribonucleic acids for the pituitary  
urix coturnix japonica).  
A:Reference number: 151241; MUID:94252550



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2001, 16:01:15 ; Search time 31.13 Seconds

(without alignments)  
234.385 Million cell updates/sec

Title: US-09-668-021-14

Perfect score: 1167

Sequence: 1 MQLSLAPCLACLIVHAFAV.....KPRPRARAKANQAELENNY 213

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

rched: 93435 seqs, 3425486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	ID	Description
1	95.5	8.2	159	1 LSHB_MEIGA	P45646 meleagris g
2	88	7.5	118	1 LSHB_PHTCA	P23330 physeler ca
3	87.5	7.5	141	1 LSHB_TRTVU	O46482 trichosurus
4	86	7.4	118	1 LSHB_BALAC	P33088 balaenopter
5	83	7.1	474	1 KLF4_MOUSE	O60793 mus musculu
6	82.5	7.1	166	1 LSHB_COTJA	P45657 coturnix co
7	81.5	7.0	470	1 KLF4_HUMAN	O44474 homo sapien
8	81	6.9	170	1 PLGF_HUMAN	P49763 homo sapien
9	80	6.9	2142	1 BAT2_HUMAN	P48634 homo sapien
10	79.5	6.8	165	1 CSHB_HUMAN	P01233 homo sapien
11	77.5	6.6	138	1 LSHB_MACRU	O46483 macropus ru
12	77.5	6.6	188	1 VEGB_MOUSE	P49766 mus musculu
13	77.5	6.6	368	1 TGF4_MOUSE	O64280 mus musculu
14	77	6.6	1480	1 SLTT_DROME	P24014 drosophila
15	76.5	6.6	301	1 UL49_HSV11	P10233 herpes simp
16	76.5	6.6	361	1 LFTB_HUMAN	O75610 homo sapien
17	75.5	6.5	128	1 LSHB_STRCA	P80664 struthio ca
18	75.5	6.5	368	1 LFTB_MOUSE	P57785 mus musculu
19	75.5	6.5	1172	1 LMB3_HUMAN	O13751 homo sapien
20	75.5	6.5	2164	1 CCAA_MOUSE	P97445 mus musculu
21	75	6.4	141	1 LSHB_CERSI	O77835 ceratotheri
22	75	6.4	180	1 DAN_HUMAN	P41271 homo sapien
23	75	6.4	355	1 KLF2_HUMAN	O9Y5W3 homo sapien
24	75	6.4	379	1 CYR6_MOUSE	P18406 mus musculu
25	75	6.4	381	1 CYR6_HUMAN	O00622 homo sapien
26	75	6.4	628	1 MSLN_HUMAN	O13421 homo sapien
27	75	6.4	951	1 SPFR8_HUMAN	O12872 homo sapien
28	74.5	6.4	322	1 YBI_MOUSE	P27817 mus musculu
29	74	6.3	174	1 BAR1_CHITE	P02849 chironomus
30	74	6.3	366	1 TGF4_HUMAN	O00292 homo sapien
31	74	6.3	563	1 MUC5_BOVIN	P98091 bos taurus
32	73.5	6.3	324	1 YBI_HUMAN	P16991 homo sapien
33	73.5	6.3	1436	1 WC11_BOVIN	P30205 bos taurus

34	73.5	6.3	3133	1 HMCT_BOMO	P98092 bombyx mori
35	73	6.3	305	1 YB3_XENLA	O00436 xenopus lae
36	73	6.3	511	1 NEK3_MOUSE	O97025 mus musculu
37	73	6.3	581	1 PRUR_BOVIN	O28172 bos taurus
38	73	6.3	644	1 VPA_BTV13	P33429 bluetongue
39	73	6.3	1955	1 AGRI_CHICK	P31696 gallus galli
40	72.5	6.2	551	1 YO11_MYCTU	P71732 mycobacteri
41	72.5	6.2	1056	1 MUC5_HUMAN	P98088 homo sapien
42	72	6.2	644	1 VPA_BTV11	P33428 bluetongue
43	72	6.2	888	1 KLTR_MOUSE	P08923 mus musculu
44	72	6.2	975	1 CDP_CANFA	P39881 canis fam11
45	72	6.2	1268	1 PCGN_MOUSE	P55066 mus musculu

## ALIGNMENTS

RESULT ID	LSHB_MEIGA	STANDARD:	PERT:	159 AA.
AC	P45646:			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	LUTROPIN BETA CHAIN PRECURSOR (LUTEINIZING HORMONE BETA SUBUNIT) (LSH-BETA) (LSH-B) (LH-B).			
DE	LHB.			
OS	Meleagris gallopavo (Common turkey).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Archosauria; Aves; Neognathae; Galliformes; Meleagridae; Meleagris.			
OX	NCBI_TaxID=9103;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Pituitary;			
RX	MEDLINE=95290073; PubMed=7772235;			
RA	You S., Foster L.K., Silsby J.L., el Halawani M.E., Foster D.N.;			
RT	"Sequence analysis of the turkey LH beta subunit and its regulation by gonadotrophin-releasing hormone and prolactin in cultured pituitary cells.";			
RT	J. Mol. Endocrinol. 14:117-129(1995).			
RL	-1- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.			
CC	-1- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN.			
CC	LUTROPIN, FOLLITROPIN AND GONADOTROPIN.			
CC	-1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).			
CC	-----			
DR	EMBL: L35519; AAA74125.1; ALT_INIT.			
DR	HSPB: P01233; IHRP.			
DR	InterPro: IPR000359; -.			
DR	InterPro: IPR001545; -.			
DR	InterPro: IPR002400; -.			
DR	Pfam: PF00007; Cys_knot; 1.			
DR	PRINTS: PR00438; GFCYSKNOT.			
DR	PROSITE: PS00261; GLYCO_HORMONE_BETA_1; 1.			
DR	PROSITE: PS00689; GLYCO_HORMONE_BETA_2; 1.			
KW	Hormone; Signal; Glycoprotein.			
FT	SIGNAL	1	39	POTENTIAL.
FT	CHAIN	40	159	LUTROPIN BETA CHAIN.
FT	DISULFID	48	96	BY SIMILARITY.
FT	DISULFID	62	111	BY SIMILARITY.
FT	DISULFID	65	149	BY SIMILARITY.
FT	DISULFID	73	127	BY SIMILARITY.
FT	DISULFID	77	129	BY SIMILARITY.

FT DISULFID 132 139 BY SIMILARITY.  
 FT CARBOHYD 52 52 N-LINKED (GLCNAC... ) (POTENTIAL).  
 SQ SEQUENCE 159 AA; 16265 MM; 52B50C8C879653C6 CRC64;

Query Match 8.2%; Score 95.5; DB 1; Length 159;  
 Best Local Similarity 27.4%; Pred. No. 0.047;  
 Matches 34; Conservative 11; Mismatches 58; Indels 21; Gaps 5;

QY 62 GGRPHHPDTRKDYSEYSCRELHYTFYVDGCRSAKPYTELVCSGCGPARLLPNAIGR 121  
 DB 43 GGRPP-----CRPINVTVAVERDECPQCAVTTTACGXCRT-----EPVYR 85  
 QY 122 VKMMRPNGPDRFCIPDRYARQVOLL-CPGGAAPRSRKRVLVASCCKRRLTRFHNSSELK 180  
 DB 86 SPGRP-PPSSCTGALRERWALMGCPITGSDPRV-LLPVALSCRCARCPITISDCTVQ 142  
 QY 181 DEGP 184  
 DB 143 GLGP 146

RESULT 2  
 ID LSHB\_PHYCA STANDARD; PRT; 118 AA.  
 AC P25330;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE LUTROPIN BETA CHAIN (LUTEINIZING HORMONE BETA SUBUNIT) (LSH-BETA)  
 GN LHB.  
 OS Physeter catodon (Sperm whale) (Physeter macrocephalus).  
 OC Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti;  
 OC Physetereidae; Physeter.  
 OX NCBI\_Taxid=9755;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=87032654; PubMed=3771098;  
 RA Pankov Y.A., Karasov V.S.;  
 RT "Primary structure of sperm whale luteinizing hormone.";  
 RL Int. J. Pept. Protein Res. 28:124-129(1986).  
 RN [2]  
 RP SEQUENCE.  
 RX MEDLINE=84281133; PubMed=6466737;  
 RA Pankov Y.A., Karasov V.S.;  
 RT "Luteinizing hormone of the sperm whale. Amino acid sequences of  
 reduced and carboxymethylated beta-subunits.";  
 RL Biochimica 49:1004-1018(1984).  
 CC -1- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING  
 THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.  
 CC -1- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA  
 CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,  
 LUTROPIN, FOLLITROPIN AND GONADOTROPIN.  
 CC -1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN  
 FAMILY.  
 CC PIR: PNCI41; PNCI41.  
 DR HSSP: P01233; IHRP.  
 DR InterPro: IPR000359; -;  
 DR InterPro: IPR001545; -;  
 DR InterPro: IPR002400; -;  
 DR Pfam: PF00007; Cys\_knot; 1.  
 DR PRINTS: PR00438; GFCYSKNOT.  
 DR PROSITE: PS00261; GLYCO\_HORMONE\_BETA\_1; 1.  
 DR PROSITE: PS00689; GLYCO\_HORMONE\_BETA\_2; 1.  
 KW Hormone; Glycoprotein.  
 FT DISULFID 9 57 BY SIMILARITY.  
 FT DISULFID 23 72 BY SIMILARITY.  
 FT DISULFID 26 110 BY SIMILARITY.  
 FT DISULFID 34 88 BY SIMILARITY.  
 FT DISULFID 38 90 BY SIMILARITY.  
 FT DISULFID 93 100 BY SIMILARITY.

FT CARBOHYD 13 13 N-LINKED (GLCNAC... )  
 SQ SEQUENCE 118 AA; 12412 MM; 81177A56382F15E7 CRC64;

Query Match 7.5%; Score 88; DB 1; Length 118;  
 Best Local Similarity 29.5%; Pred. No. 0.17;  
 Matches 36; Conservative 13; Mismatches 55; Indels 18; Gaps 6;

QY 80 CRELHYTFEYVDGCRSAKPYTELVCSGQC-GPARLLPNAIGRKVMRPNPDRFCIPDR 138  
 DB 9 CRPINATLAONZACPVICITFTTSICAGYCPSMVRVLPAL-----PPVZPVCTYRQ 61  
 QY 139 YRAGRVOOLL-CPGGAAPR-SRKRVLVASCCKRRLTRFHNSSELDGPEFARPOK-GRKP 195  
 DB 62 LRFASIRLPDGPPEVNPVNSFPVALSCHCGCRLLS-----SDCGRGAQPLACNRSR 114  
 QY 196 RP 197  
 DB 115 RP 116

RESULT 3  
 ID LSHB\_TRIUV STANDARD; PRT; 141 AA.  
 AC O46482;  
 DT 01-OCT-2000 (Rel. 40, Created)  
 DT 01-OCT-2000 (Rel. 40, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE LUTROPIN BETA CHAIN PRECURSOR (LUTEINIZING HORMONE BETA SUBUNIT) (LSH-BETA) (LSH-B).  
 GN LHB.  
 OS Trichosurus vulpecula (Brush-tailed possum).  
 OC Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;  
 OC Mammalia; Metatheria; Diprotodontia; Phalangeridae; Trichosurus.  
 OX NCBI\_Taxid=9337;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=pituitary;  
 RX MEDLINE=98345424; PubMed=9680384;  
 RA Harrison G.A., Deane E.M., Cooper D.W.;  
 RT "cDNA cloning of luteinizing hormone subunits from brushtail possum  
 and red kangaroo.";  
 RL Mamm. Genome 9:638-642(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Lawrence S.B., McNatty K.P., Fidler A.E.;  
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING  
 THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.  
 CC -1- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA  
 CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,  
 LUTROPIN, FOLLITROPIN AND GONADOTROPIN.  
 CC -1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN  
 FAMILY.  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-stb.ch/announce/>  
 or send an email to [license@isb-stb.ch](mailto:license@isb-stb.ch)).  
 CC EMBL: AF017448; AAC96019.1; -;  
 DR EMBL: AF090388; AAC63526.1; -;  
 DR HSSP: P01233; IHRP.  
 DR InterPro: IPR000359; -;  
 DR InterPro: IPR001545; -;  
 DR Pfam: PF00007; Cys\_knot; 1.  
 DR PROSITE: PS00261; GLYCO\_HORMONE\_BETA\_1; 1.  
 DR PROSITE: PS00689; GLYCO\_HORMONE\_BETA\_2; 1.  
 KW Hormone; Signal; Glycoprotein.  
 FT SIGNAL 1 22 BY SIMILARITY.

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2001, 16:00:36 : Search time 89.34 Seconds

(without alignments)  
315.435 Million cell updates/sec

Title: US-09-668-021-14

Perfect score: 1167  
Sequence: 1 MGLSLAPCLACLVHAFVA.....KPRPRARGAKANOAELENAY 213

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

rchd: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: SPREMBL\_16:\*  
2: sp\_archaea:\*  
3: sp\_bacteria:\*  
4: sp\_fungi:\*  
5: sp\_human:\*  
6: sp\_invertebrate:\*  
7: sp\_mammal:\*  
8: sp\_mhc:\*  
9: sp\_organelle:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_unclassified:\*  
13: sp\_vertebrate:\*  
14: sp\_virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	256.5	22.0	134	4 Q9Y3U3	Q9Y3U3 homo sapien
2	96	8.2	272	13 Q9PWB0	Q9PWB0 gallus gall
3	94	8.1	272	13 Q9PUK2	Q9PUK2 gallus gall
4	93.5	8.0	1114	11 Q9JRW7	Q9JRW7 mus musculu
5	92	7.9	712	4 Q9Y557	Q9Y557 homo sapien
6	92	7.9	720	4 Q9NY23	Q9NY23 homo sapien
7	90.5	7.8	493	4 Q9UEW3	Q9UEW3 homo sapien
8	88	7.5	168	4 Q9H772	Q9H772 homo sapien
9	87	7.5	184	13 Q73755	Q73755 gallus gall
10	87	7.5	381	11 Q88995	Q88995 mus musculu
11	87	7.5	988	6 Q97867	Q97867 sus scrofa
12	87	7.5	1537	11 Q9WTW6	Q9WTW6 mus musculu
13	87	7.5	1581	11 Q9ROB6	Q9ROB6 mus musculu
14	86.5	7.4	493	4 Q9S376	Q9S376 homo sapien
15	86	7.4	925	2 Q9F4J0	Q9F4J0 porphyromon
16	85.5	7.3	481	4 Q9HBZ6	Q9HBZ6 homo sapien
17	85.5	7.3	791	11 Q70368	Q70368 rattus norv
18	83	7.1	168	11 Q88273	Q88273 mus musculu
19	83	7.1	483	11 Q9R255	Q9R255 mus musculu

20	82.5	7.1	184	4 Q60565	Q60565 homo sapien
21	82	7.0	1035	5 Q9NEG1	Q9NEG1 drosophila
22	81.5	7.0	470	4 Q9UNP3	Q9UNP3 homo sapien
23	81	6.9	2157	4 Q95875	Q95875 homo sapien
24	81	6.9	3163	2 Q92373	Q92373 yersinia pe
25	80.5	6.9	215	8 Q9WF86	Q9WF86 beta vulgar
26	80.5	6.9	1081	4 Q76065	Q76065 homo sapien
27	80.5	6.9	1704	5 Q94446	Q94446 chironomus
28	80	6.9	379	11 Q9ES72	Q9ES72 rattus norv
29	80	6.9	1096	10 Q9XE24	Q9XE24 oryza sativ
30	80	6.9	3161	2 Q54511	Q54511 yersinia en
31	78.5	6.7	492	11 Q9Z1K6	Q9Z1K6 mus musculu
32	78.5	6.7	761	3 Q9HEA3	Q9HEA3 neurospora
33	78.5	6.7	2098	5 Q25757	Q25757 plasmodium
34	78.5	6.7	2114	5 Q97267	Q97267 plasmodium
35	78	6.7	343	11 Q61406	Q61406 mus musculu
36	77.5	6.6	184	11 Q35793	Q35793 rattus norv
37	77.5	6.6	184	11 Q70326	Q70326 mus musculu
38	77.5	6.6	270	13 P70041	P70041 xenopus lae
39	77.5	6.6	482	14 Q71654	Q71654 human immun
40	77.5	6.6	1637	6 Q9XSV8	Q9XSV8 bos taurus
41	77	6.6	267	4 Q95813	Q95813 homo sapien
42	77	6.6	355	4 Q9UKR6	Q9UKR6 homo sapien
43	77	6.6	384	10 Q9PSQ3	Q9PSQ3 oryza sativ
44	77	6.6	385	5 Q27008	Q27008 toxoplasma
45	77	6.6	826	10 Q9FU12	Q9FU12 oryza sativ

## ALIGNMENTS

RESULT	1	PRELIMINARY:	PRT:	134 AA.
Q9Y3U3	Q9Y3U3			
AC	Q9Y3U3:			
DT	01-NOV-1999 (TREMBLrel. 12, Created)			
DT	01-NOV-1999 (TREMBLrel. 12, Last sequence update)			
DT	01-MAY-2000 (TREMBLrel. 13, Last annotation update)			
DE	HYPOHETICAL 15.3 KDA PROTEIN (FRAGMENT).			
GN	DKEZP564D206.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=BRAIN;			
RA	Wambolt R., Heubner D., Mewes H.W., Gassenhuber J., Wiemann S.;			
RL	Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AL050024; CAB43243.1; -			
DR	InterPro: IPR000359; -			
DR	PROSITE: PS01225; CTCK_2; 1.			
KW	Hypothetical protein.			
FT	NON_TER			
SQ	SEQUENCE	134 AA; 15324 MW; A0B7A8650D2EB6F1 CRC64;		

Query Match 22.0%; Score 256.5; DB 4; Length 134;  
Best Local Similarity 41.9%; Pred. No. 1.8e-18;  
Matches 52; Conservative 24; Mismatches 43; Indels 5; Gaps 3;

QY	80 CRELHYTRFVYDGPFRSAKPVTELVCSGCGPARLLPNAIG--RVKWM-RPNGPDFRCI 135	
DB	3 CRELRSKRYISDGGCTSIPLKELVCAEGEDLLVPVLPNWIGGVTGRYMSRRSSQEMRCV 62	
QY	136 PDRRAQVOLLGCGAAPSRAKRYLVASCKRKLRTFHHQSLKDPGPTARPGGRFP 195	
DB	63 NDKRTRTRIQLQCODGST-RYKLTIVYACKCKRYTRQHNESHNSPAPKPVQHHRE 121	
QY	196 RPRA 199	
DB	122 RKRA 125	

```

RESULT 2
ID 09PWB0 PRELIMINARY; PRT; 272 AA.
AC 09PWB0;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE CERBERUS HOMOLOG.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9439862; PubMed=10508582;
RA Zhu L., Marvin M.J., Gardiner A., Lassar A.B., Mercola M., Stern C.D.,
RA Levin M.;
RT "Cerberus regulates left-right asymmetry of the embryonic head and
heart."
CURR. BIOL. 9:931-938(1999).
EMBL: AF139721; AAD51610.1; -.
InterPro: IPR000359; -.
DR PROSITE: PS01225; CTCK_2; 1.
DR SMART: SM00041; CT; 1.
SQ SEQUENCE 272 AA; 31224 MW; 36EAC9F719711BCA CRC64;

Query Match 8.2%; Score 96; DB 13; Length 272;
Best Local Similarity 22.2%; Pred. No. 0.065;
Matches 30; Conservative 22; Mismatches 57; Indels 26; Gaps 5;

OY 45 EPPQELN-----NOTMNAENGRRPHHPYDTKQVSEYSCRELHYTFYDGCPSAKPV 100
DB 129 EPPYKDKAKFTWHDHMLKRNKNSAEVVLPIKTNEHQCRTLPESQVAHSCSEKV-IV 187
OY 101 TELVSGGCGPARLLPNAIGRVKWRPNRNGPDR-----CIPDRYRAQVOLLCPGAA 153
DB 188 QNNLCFGKCSS-----FHVGPDDRLLYTFCSKCLPTKFSMKHFDLNCI-SSV 233
OY 154 PRSRKRVLVASCCK 168
DB 234 PVVKKVMIVEECNCE 248

RESULT 3
O9PUK2 PRELIMINARY; PRT; 272 AA.
O9PUK2;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE CARONTE.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Rodriguez Esteban C., Capdevila J., Economides A.N., Pascual J.,
RA Ortiz A., Izpisua Belmonte J.C.;
RT "Caronte, a novel cer-like protein, mediates the establishment of
RT embryonic left-right asymmetry."
RL Nature 0:0-0(1999).
EMBL: AF179484; AAD55581.1; -.
InterPro: IPR000359; -.
DR InterPro: IPR001839; -.
DR ProDom: PD000357; -.
DR PROSITE: PS01225; CTCK_2; 1.

```

```

DR SMART: SM00041; CT; 1.
SQ SEQUENCE 272 AA; 31201 MW; 0D89729715771BC2 CRC64;

Query Match 8.1%; Score 94; DB 13; Length 272;
Best Local Similarity 22.2%; Pred. No. 0.1;
Matches 30; Conservative 21; Mismatches 58; Indels 26; Gaps 5;

OY 45 EPPQELN-----NOTMNAENGRRPHHPYDTKQVSEYSCRELHYTFYDGCPSAKPV 100
DB 129 EPPYKDKAKFTWHDHMLKRNKNSAEVVLPIKTNEHQCRTLPESQVAHSCSEKV-IV 187
OY 101 TELVSGGCGPARLLPNAIGRVKWRPNRNGPDR-----CIPDRYRAQVOLLCPGAA 153
DB 188 QNNLCFGKCSS-----FHVGPDDRLLYTFCSKCLPTKFSMKHFDLNCI-SSV 233
OY 154 PRSRKRVLVASCCK 168
DB 234 PVVKKVMIVEECNCE 248

RESULT 4
O9UKW7 PRELIMINARY; PRT; 1114 AA.
AC 09UKW7;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE CYSTEINE-RICH PROTEIN NFX-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Ariotta P., Jenkins N.A., Jay G., Ono S.J.;
RT "Isolation of a Full-Length Murine NFX-1 cDNA, Its Chromosomal
RT Location and Developmental Expression."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF223576; AAF34700.1; -.
DR InterPro: IPR000967; -.
DR InterPro: IPR001374; -.
DR InterPro: IPR001841; -.
DR InterPro: IPR001965; -.
DR Pfam: PF01422; zf-NF-XL; 8.
DR Pfam: PF01424; R3H; 1.
SQ SEQUENCE 1114 AA; 123877 MW; 857C90260D28EAB CRC64;

Query Match 8.0%; Score 93.5; DB 11; Length 1114;
Best Local Similarity 23.6%; Pred. No. 0.5;
Matches 48; Conservative 26; Mismatches 90; Indels 39; Gaps 9;

OY 21 VESGQMOAFKND-----ATEIIRGLREYDEPPQELENNQTMNRAENGRR-PPHHPTDKDV 75
DB 98 LONOPWQAKRNEKHONRNKKAQGLSEQTSSTSSLE---SVARSSEGTNPRESHPSSE-- 152
OY 76 SEYSCRELHYTFYDGCPSAKPVTELVSGGCGPARLLPNAIGRVK--WRRPNRNGPDR 133
DB 153 -----KEVVIADPRGAKPKKAAQLYNNRG--PKAKGRARSMGMRMSKSE 198
OY 134 C-IPDRYRAQVOLLCPGGAAPRSRKRVLVASCCKRLTFPHNOSLUKDFGPETARPQKG 192
DB 199 DEIDPPMWFPTLTQIASCRKP-----VDDPCVCRRNEORRYRQKRPLGSGRARPRG 252
OY 193 RKP-----RRRARGAKANQAEI 209
DB 253 RNPKEQSQRHINAGPKTNMSPI 275

RESULT 5
O9Y557 PRELIMINARY; PRT; 712 AA.
ID 09Y557

```

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2001, 15:48:43 ; Search time 80.28 Seconds

(without alignments)  
160.848 Million cell updates/sec

Title: US-09-668-021-14

Perfect score: 1167  
Sequence: 1 MQLSLAPCLACLLVNAAFVA.....KPRRARGAKANQAELENNY 213

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

ched: 412676 segs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A.GeneSeq 0601:\*

1: /SIDS8/gcgdata/geneSeq/AAl980.DAT:\*  
2: /SIDS8/gcgdata/geneSeq/AAl981.DAT:\*  
3: /SIDS8/gcgdata/geneSeq/AAl982.DAT:\*  
4: /SIDS8/gcgdata/geneSeq/AAl983.DAT:\*  
5: /SIDS8/gcgdata/geneSeq/AAl984.DAT:\*  
6: /SIDS8/gcgdata/geneSeq/AAl985.DAT:\*  
7: /SIDS8/gcgdata/geneSeq/AAl986.DAT:\*  
8: /SIDS8/gcgdata/geneSeq/AAl987.DAT:\*  
9: /SIDS8/gcgdata/geneSeq/AAl988.DAT:\*  
10: /SIDS8/gcgdata/geneSeq/AAl989.DAT:\*  
11: /SIDS8/gcgdata/geneSeq/AAl990.DAT:\*  
12: /SIDS8/gcgdata/geneSeq/AAl991.DAT:\*  
13: /SIDS8/gcgdata/geneSeq/AAl992.DAT:\*  
14: /SIDS8/gcgdata/geneSeq/AAl993.DAT:\*  
15: /SIDS8/gcgdata/geneSeq/AAl994.DAT:\*  
16: /SIDS8/gcgdata/geneSeq/AAl995.DAT:\*  
17: /SIDS8/gcgdata/geneSeq/AAl996.DAT:\*  
18: /SIDS8/gcgdata/geneSeq/AAl997.DAT:\*  
19: /SIDS8/gcgdata/geneSeq/AAl998.DAT:\*  
20: /SIDS8/gcgdata/geneSeq/AAl999.DAT:\*  
21: /SIDS8/gcgdata/geneSeq/AAl2000.DAT:\*  
22: /SIDS8/gcgdata/geneSeq/AAl2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1167	100.0	213	21	AAV96433
2	1094	93.7	211	21	AAV96432
3	1073	91.9	213	21	AAV96431
4	1050	90.0	213	21	AAV96429
5	1050	90.0	213	21	AAV96429
6	1050	90.0	213	21	AAV96429
7	1041	89.2	213	21	AAV96436
8	1040	89.1	213	21	AAV96430
9	995	81.8	367	21	AAV964105
10	861.5	73.8	176	21	AAV96434
11	343.5	29.4	206	21	AAV75981

12	343.5	29.4	206	21	AAV76031	Murine skin cell p
13	343.5	29.4	206	22	AAV55920	Skin cell protein,
14	343.5	29.4	206	22	AAV55970	skin cell protein,
15	340	29.1	206	18	AAV09408	Human small CCN-11
16	340	29.1	206	19	AAV58704	Human small CCN-11
17	340	29.1	206	20	AAV95721	Homo sapiens fetal
18	340	29.1	206	21	AAV10233	Human adult retina
19	201	17.2	50	20	AAV12009	Human 5' EST seque
20	115	9.9	116	18	AAV27654	Secreted protein A
21	115	9.9	116	18	AAV44090	Human secreted pro
22	98.5	8.4	102	20	AAV03225	Human acid sequenc
23	98.5	8.4	102	21	AAV10277	Human fetal kidney
24	94.5	8.1	252	21	AAV57316	Alpha-mating facto
25	92	7.9	712	21	AAV43147	Human ORFX ORF2911
26	88	7.5	168	21	AAV84014	Amino acid sequenc
27	87	7.5	184	19	AAV29727	DNA and b57 protel
28	86.5	7.4	165	20	AAV99515	Glycoprotein hormo
29	86	7.4	921	20	AAV34522	Porphyromonas ging
30	86	7.4	922	20	AAV34522	Porphyromonas ging
31	86	7.4	925	20	AAV34520	Porphyromonas ging
32	86	7.4	938	20	AAV34392	Porphyromonas ging
33	85.5	7.3	145	12	AAV5109	hCG/bhLH chimera, E
34	85.5	7.3	175	21	AAV57315	Human betahCG/beta
35	84.5	7.2	229	9	AAV82071	pBoel1743-2-9-8 in
36	83.5	7.2	145	12	AAV5125	hCG/hhLH chimera, A
37	83.5	7.2	145	21	AAV87479	Human chorionic go
38	83.5	7.2	158	20	AAV43266	Human chorionic go
39	83.5	7.2	204	20	AAV43265	Human CG beta subu
40	83	7.1	168	21	AAV51132	Murine cerebral ne
41	83	7.1	474	22	AAV31245	Amino acid sequenc
42	82.5	7.1	145	12	AAV5099	hCG/bhLH chimera, D
43	82.5	7.1	184	19	AAV69293	Human b57 protein
44	82.5	7.1	184	20	AAV42173	Mouse DN protein
45	82.5	7.1	184	21	AAV95961	Human gremlin. Ho

#### ALIGNMENTS

RESULT 1	AAV96433	standard: Protein: 213 AA.
XX	AAV96433	
AC	AAV96433	
XX		
DT	12-SEP-2000 (first entry)	
XX		
DE	Rat TGF-beta binding protein (BEER).	
XX		
KW	osteopathic: transforming growth factor-beta: TGF-beta: binding protein;	
KW	BEER, gene therapy: antisense therapy; fracture; bone mineralization.	
OS	Rattus norvegicus.	
XX		
PN	WO200032773-A1.	
XX		
PD	08-JUN-2000.	
XX		
PF	24-NOV-1999; 99WO-US27990.	
XX		
FR	27-NOV-1998; 98US-0110283.	
XX		
PA	(DARW-) DARWIN DISCOVERY LTD.	
XX		
PI	Brunkow ME, Gales DJ, Kovacevich B, Mulligan JT, Paepker BW;	
XX	Van Ness J, Winkler DG;	
XX		
DR	WPI: 2000-412321/35.	
XX	N-PSDB; AAA29059.	
XX		
PT	Nucleic acids (I) encoding a transforming growth factor beta binding	
PT	protein, useful for identifying agents for treating osteopenia,	
PT	osteoporosis and fractures	

XX Claim 6; Page 125-126; 162pp; English.  
PS  
XX This shows a rat transforming growth factor-beta (TGF-beta) binding  
CC protein designated rBEER. The cDNA and protein may be used for  
CC prevention, treatment and diagnosis of diseases associated with  
CC inappropriate BEER expression. For example, they may be used to treat  
CC disorders associated with decreased TGF-beta BP expression. The cDNA or  
CC vectors may be administered to treat diseases by rectifying mutations or  
CC deletions in a patient's genome that affect the activity of BEER by  
CC expressing inactive proteins or to supplement the patients own production  
CC of BEER polypeptides. The nucleic acids may be used for recombinant  
CC production of BEER, gene therapy, antisense therapy, as probes for  
CC diagnostic assays and for functional studies. BEER may be used to raise  
CC antibodies and for identification of BEER modulators. BEER antagonists  
CC may be used to increase bone mineral content for the treatment of  
CC disorders such as osteopenia, osteoporosis, fractures and other disorders  
CC associated with low mineral content.  
XX  
SQ Sequence 213 AA;  
100.0%; Score 1167; DB 21; Length 213;  
Best Local Similarity 100.0%; Pred. No. 6,7e-102;  
Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Query Match  
1 MQLSLAPCLACLLVHAFAVAVESGQWAFKNDATETIIPGLREYEPPEOLENNQTNRAE 60  
Db 1 mqlslapclacilivhaafavessqgwafkndateilipglreyeppeqlennqtnrae 60  
QY 61 NGRPRPHHYDTRDSEYSCRELHAYTRFVTGDCPSAKPVTIELVCSGCGPARLPPNAIG 120  
Db 61 ngrprphhydtkdvseycrelhytrfvtgdcpsakpvtelvcsgcgparlppnaig 120  
QY 121 RVKWMRPNGPDRFCIPDRYRAQROVOLLCPGAAPRSKRVLVASCKKRLTRFHNSELK 180  
Db 121 rvkwmrpnpgdrfcipdryraqrvqlcpggaaprskrvllvasckckrltrfhngselk 180  
QY 181 DFGPETARPOKGRKPRPARAGAKANOAELENAY 213  
Db 181 dfgpetarpqkgrkprparagakanagaelenay 213  
RESULT 2  
AA96432 standard; Protein; 211 AA.  
XX  
AC AAY96432;  
XX  
DT 12-SEP-2000 (first entry)  
XX  
XX Murine TGF-beta binding protein (BEER).  
XX  
KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;  
KW BEER; gene therapy; antisense therapy; fracture; bone mineralization.  
XX  
OS Mus musculus.  
XX  
PN MO200032773-A1.  
XX  
PD 08-JUN-2000.  
XX  
PF 24-NOV-1999; 99WO-US27990.  
XX  
PR 27-NOV-1998; 98US-0110283.  
XX  
PA (DARW-) DARWIN DISCOVERY LTD.  
XX  
PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paeper BW;  
PI Van Ness J, Winkler DG;  
XX  
XX WPI: 2000-412321/35.  
XX  
DR N-PSDB; AAA29058.

XX Nucleic acids (I) encoding a transforming growth factor beta binding  
PT protein, useful for identifying agents for treating osteopenia,  
PT osteoporosis and fractures  
XX  
PS Claim 5; Page 124; 162pp; English.  
XX  
XX This shows a murine transforming growth factor-beta (TGF-beta)  
CC binding protein designated mBEER. The cDNA and protein may be used for  
CC prevention, treatment and diagnosis of diseases associated with  
CC inappropriate BEER expression. For example, they may be used to treat  
CC disorders associated with decreased TGF-beta BP expression. The cDNA or  
CC vectors may be administered to treat diseases by rectifying mutations or  
CC deletions in a patient's genome that affect the activity of BEER by  
CC expressing inactive proteins or to supplement the patients own production  
CC of BEER polypeptides. The nucleic acids may be used for recombinant  
CC production of BEER, gene therapy, antisense therapy, as probes for  
CC diagnostic assays and for functional studies. BEER may be used to raise  
CC antibodies and for identification of BEER modulators. BEER antagonists  
CC may be used to increase bone mineral content for the treatment of  
CC disorders such as osteopenia, osteoporosis, fractures and other disorders  
CC associated with low mineral content.  
XX  
SQ Sequence 211 AA;  
93.7%; Score 1094; DB 21; Length 211;  
Best Local Similarity 93.9%; Pred. No. 4,7e-95;  
Matches 200; Conservative 3; Mismatches 6; Indels 2; Gaps 1;  
Query Match  
1 MQLSLAPCLACLLVHAFAVAVESGQWAFKNDATETIIPGLREYEPPEOLENNQTNRAE 60  
Db 1 mqlslapclacilivhaafavegqgwafkndatevlpglreyeppep--ennqtnrae 58  
QY 61 NGRPRPHHYDTRDSEYSCRELHAYTRFVTGDCPSAKPVTIELVCSGCGPARLPPNAIG 120  
Db 59 ngrprphhydtkdvseycrelhytrfvtgdcpsakpvtelvcsgcgparlppnaig 118  
QY 121 RVKWMRPNGPDRFCIPDRYRAQROVOLLCPGAAPRSKRVLVASCKKRLTRFHNSELK 180  
Db 119 rvkwmrpnpgdrfcipdryraqrvqlcpggaaprskrvllvasckckrltrfhngselk 178  
QY 181 DFGPETARPOKGRKPRPARAGAKANOAELENAY 213  
Db 179 dfgpetarpqkgrkprparagakanagaelenay 211  
RESULT 3  
AA96431 standard; Protein; 213 AA.  
XX  
AC AAY96431;  
XX  
DT 12-SEP-2000 (first entry)  
XX  
XX Vervet TGF-beta binding protein (BEER).  
XX  
DE Vervet TGF-beta binding protein (BEER).  
XX  
KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;  
KW BEER; gene therapy; antisense therapy; fracture; bone mineralization.  
XX  
OS Cercopithecus pygerythrus.  
XX  
PN MO200032773-A1.  
XX  
PD 08-JUN-2000.  
XX  
PF 24-NOV-1999; 99WO-US27990.  
XX  
PR 27-NOV-1998; 98US-0110283.  
XX  
PA (DARW-) DARWIN DISCOVERY LTD.  
XX  
PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paeper BW;



```

?
? APPLICANT: Watson, James D.
? APPLICANT: Strachan, Lorna
? APPLICANT: Sleeman, Matthew
? APPLICANT: Onrust, Rene
? APPLICANT: Murison, James Greg
? TITLE OF INVENTION: Compositions Isolated From Skin Cells
? TITLE OF INVENTION: and Methods For Their Use
? FILE REFERENCE: 11000.101c1
? CURRENT APPLICATION NUMBER: US/09/188,930A
? CURRENT FILING DATE: 1998-11-09
? NUMBER OF SEQ ID NOS: 348
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO 286
? LENGTH: 206
? TYPE: PRT
? ORGANISM: Mouse
?
US-09-188-930-286

```

Query Match	29.4%	Score 343.5;	DB 4;	Length 206;
Best Local Similarity	39.3%	Pred. No. 1.4e-28;		
Matches 81;	Conservative 32;	Mismatches 72;	Indels 21;	Gaps 8

[illegible]

US-RSULT 3  
US-088-847B-2  
Sequence 2, Application US/08468847B  
Patent No. 5780263  
GENERAL INFORMATION:  
APPLICANT: Hastings, Gregg A. and Adams, Mark D.  
TITLE OF INVENTION: Human CCN-Like Growth Factor  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,  
ADDRESS: CECCHI, STEWART & OLSTEIN  
STREET: 6 BECKER FARM ROAD  
CITY: ROSELAND  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 INCH DISKETTE  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/468,847B  
FILING DATE: 6 June 1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: MULLINS, J.G.  
REGISTRATION NUMBER: 33,073  
REFERENCE/DOCKET NUMBER: 325800-442  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 206 AMINO ACIDS  
TYPE: AMINO ACID  
STRANDEDNESS:  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PROTEIN

US-08-468-847B-2

Query Match	29.1%	Score 340;	DB 1;	Length 206;
Best Local Similarity	39.3%	Pred. NO. 3.4e-28;		
Matches 77; Conservative	33;	Mismatches 70;	Indels 16;	Gaps 6;

QY 9 LACILVHAFAVAESOGQAKRNATATELIPOLREYPEPEOLENNONMRAENGGR-PH 6  
 11 : : : : : 11111111 : : : : : 1111 :  
 Db 13 LACILMKSCLE-----AKRNATELIVSHYVKPVPAMH-SSNLTLNQARNGRGFSN 62  
 QY 68 HPYDTRDVSEYSCRELAHTRFVTVDPCRSKAPVTELVCSGCGGPARLLPNAIG---RVKM 124  
 11 : : : : : 1111 : : : : : 1111111 : : : : : 11 :  
 Db 63 TGLDENTRVYQGCERLSTRTKYISDGCSTISPLKELVCAGECLEPLPLPMWIGGGYGTKY 122  
 QY 125 W-RPENGPFRCILPDRIYRAQRIQLLCPGGAAPRSKVKALVASCKCKRLITRFHNOSELKDFG 163  
 11 : : : : : 1111 : : : : : 11 : : : : : 1111 : : : : : 11 :  
 Db 123 WSRSSQSCMRFCVNMKTRQRIQLCCODGST-RYKVIIVYVCAKCKCKRYTRHNESSHNEFS 161  
 QY 184 PETAPQKGRKRPRA 199  
 11 : : : : : 11 : : : : : 11 : : : : : 11 : : : : : 11 :  
 Db 182 MSPAKPVQHHREKRKA 197

US-08-4478-20  
4  
Sequence 20, Application US/08468847B  
Patent No. 5780263  
GENERAL INFORMATION:  
APPLICANT: Haslins, Gregg A. and Adams, Mark D.  
TITLE OF INVENTION: Human CCN-Like Growth Factor  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,  
ADDRESSEE: CECCHI, STEWART & OLSTEIN  
STREET: 6 BECKER FARM ROAD  
CITY: ROSELAND  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 INCH DISKETTE  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/468,847B  
FILING DATE: 6 June 1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: MULLINS, J.G.  
REGISTRATION NUMBER: 33,073  
REFERENCE/DOCKET NUMBER: 325800-442  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 206 AMINO ACIDS  
TYPE: AMINO ACID  
STRANDEDNESS:





C:Superfamily: pituitary glycoprotein hormone beta chain

Query Match 10.3%; Score 99.5; DB 2; Length 159;

Best Local Similarity 27.4%; Pred. No. 0.042; Mismatches 58; Indels 21; Gaps 5;

Matches 34; Conservative 11; Mismatches 58; Indels 21; Gaps 5;

QY 31 GGRPHHPEETKDAEYSCRELHFTRYVTDGPCRSAPKPYTELVCSGGCGPARLLPNAIGR 90

Db 43 GGRPP-----CRPINVTVAVEKDECPQCMATTTACGTCRTREPV-----EPYR 85

QY 91 GKMMRRSGDFRCIPDRYRAQVOLL-CPGGAAPRAKRVRLVASCCKRLTRFHNSFLK 149

Db 86 SPGGRP--POSSCTYCALRYERMLMGCPISDPRV-LLPVALLSCRCARCPATSDCTVQ 142

QY 150 DFGR 153

Db 143 GLGP 146

RESULT 3

1091  
Luteinizing hormone precursor - chicken

C:Alternate names: luteinizing hormone beta chain

C:Species: Gallus gallus (chicken)

C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 08-Dec-1995

C:Accession: A61091

R:Node: T.; Ando, H.; Ueda, T.; Kubokawa, K.; Higashinakagawa, T.; Ishii, S.

J. Mol. Endocrinol. 3, 129-137, 1989

A>Title: Molecular cloning and nucleotide sequence analysis of the putative cDNA for the

A:Reference number: A61091; MUID:89374710

A:Accession: A61091

A>Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-158 <NOC>

C:Superfamily: pituitary glycoprotein hormone beta chain

C:Keywords: glycoprotein; hormone; pituitary

F:1-33/Domain: signal sequence #status predicted <SIG>

F:48-73,62-96,65-127,77-149,111-139,129-132/Disulfide bonds: #status predicted

F:52/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.8%; Score 94.5; DB 2; Length 158;

Best Local Similarity 26.2%; Pred. No. 0.12; Mismatches 56; Indels 25; Gaps 5;

Matches 33; Conservative 12; Mismatches 56; Indels 25; Gaps 5;

QY 31 GGRPHHPEETKDAEYSCRELHFTRYVTDGPCRSAPKPYTELVCSGGCGPARLLPNAIGR 90

Db 43 GGRPP-----CRPINVTVAVEKDECPQCMATTTACGTCRTREPV----- 83

QY 91 GKMMRRSG--PDFRCIPDRYRAQVOLL-CPGGAAPRAKRVRLVASCCKRLTRFHNSFLK 147

Db 84 --YRSPGLGPPQSCCTYGALRYERMLMGCPISDPRV-LLPVALLSCRCARCPATSDCT 140

QY 148 LKDFGR 153

Db 141 VQGLGP 146

RESULT 4

T22025  
hypothetical protein F40E10.4 - Caenorhabditis elegans (fragment)

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000

C:Accession: T22025

R:Smyle, R.

submitted to the EMBL Data Library, February 1996

A:Reference number: Z19503

A:Accession: T22025

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-601 <MUI>

A:Cross-references: EMBL:Z69792; PIDN:CA93668.1; GSPDB:GN00028; CESP:F40E10.4

A:Experimental source: clone F40E10

C:Genetics:

A:Gene: CESP:F40E10.4

A:Map position: X

Query Match 9.3%; Score 89.5; DB 2; Length 601;

Best Local Similarity 20.6%; Pred. No. 1.3; Mismatches 45; Indels 19; Gaps 2;

Matches 21; Conservative 17; Mismatches 45; Indels 19; Gaps 2;

QY 36 HHPETKDAEYSCRELHFTRYVTDGPCRSAPKPYTELVCSGGCGPARLLPNAIGR 95

Db 510 HESGHCDEKRIKCKOKRRRHIECKSVDRITACNGYCG----- 553

QY 96 PSGBPFRFCIPDRYRAQVOLL-CPGGAAPRAKRVRLVASCCK 137

Db 554 --GEONCTAVKKKORRVKMKNGT--KISTVHIIRCOCE 592

RESULT 5

I51242  
Luteinizing hormone beta-subunit - quail

C:Species: Coturnix coturnix (quail)

C>Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 16-Jul-1999

C:Accession: I51242

R:Ando, H.; Ishii, S.

Gen. Comp. Endocrinol. 93, 357-368, 1994

A>Title: Molecular cloning of complementary deoxyribonucleic acids for the pituitary

A:Reference number: I51241; MUID:94252550

A:Accession: I51242

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-166 <AND>

A:Cross-references: GB:S70834; NID:9546921; PIDN:AA830867.1; PID:9546922

C:Superfamily: pituitary glycoprotein hormone beta chain

Query Match 8.6%; Score 83.5; DB 2; Length 166;

Best Local Similarity 24.8%; Pred. No. 1.3; Mismatches 56; Indels 25; Gaps 5;

Matches 31; Conservative 13; Mismatches 56; Indels 25; Gaps 5;

QY 32 GGRPHHPEETKDAEYSCRELHFTRYVTDGPCRSAPKPYTELVCSGGCGPARLLPNAIGR 91

Db 52 GGRPP-----CRPINVTVAVEKDECPQCMATTTACGTCRTREPV----- 91

QY 92 KMMRRSG--PDFRCIPDRYRAQVOLL-CPGGAAPRAKRVRLVASCCKRLTRFHNSFLK 148

Db 92 -YRSPGLGPPQSCCTYGALRYERMLMGCPISDPRV-LLPVALLSCRCARCPATSDCTV 149

QY 149 KDFGR 153

Db 150 QGLGP 154

RESULT 6

A71251  
DNA repair protein sms - syphilis spirochete

C:Alternate names: DNA repair protein radx

C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)

C>Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 19-Jan-2001

C:Accession: A71251

R:Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, R.; G

erson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Ulterback, T.; M

they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.

Science 281, 375-388, 1998

A>Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.

A:Reference number: A71251

A:Accession: A71251

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-455 <COI>

A:Cross-references: GB:AE001270; GB:AE000520; NID:93323350; PIDN:AAC65973.1; PID:93332

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: November 7, 2001, 16:01:15 ; Search time 31.13 Seconds  
(without alignments)  
193.671 Million cell updates/sec

Title: US-09-668-021-16

Perfect score: 967  
Sequence: 1 NDATETIPELGEYEPPELPEL.....RPQGRKLPRANGTKASRA 176

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

riched: 93435 segs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	99.5	10.3	159	1	LSHB_MEIGA
2	83.5	8.6	166	1	LSHB_COTJA
3	82.5	8.5	141	1	LSHB_TRIVU
4	81	8.4	455	1	RADA_TREPA
5	80	8.3	118	1	LSHB_PHYCA
6	79.5	8.2	1480	1	SLIT_DROME
7	79	8.2	470	1	KLFA_HUMAN
8	79	8.2	682	1	SNK_MOUSE
9	79	8.2	682	1	SNK_RAT
10	79	8.2	685	1	SNK_HUMAN
11	78	8.1	118	1	LSHB_BAIAC
12	77.5	8.0	128	1	LSHB_STRCA
13	77	8.0	243	1	TOMB_ENTAE
14	76.5	7.9	628	1	MSLN_HUMAN
15	75.5	7.8	141	1	LSHB_CERST
16	75.5	7.8	474	1	KLFA_MOUSE
17	75	7.8	165	1	CGHB_PAPAN
18	74	7.7	398	1	MOBI_XENLA
19	74	7.7	544	1	H15_DROME
20	73.5	7.6	870	1	BCAL_HUMAN
21	73	7.5	138	1	LSHB_MACRU
22	73	7.5	256	1	OBFL_HUMAN
23	72	7.4	143	1	LSHB_FELCA
24	72	7.4	165	1	CGHB_HUMAN
25	71.5	7.4	450	1	A2MA_MOUSE
26	71.5	7.3	1134	1	TIE1_MOUSE
27	70.5	7.3	1496	1	CA25_HUMAN
28	70	7.2	178	1	DAN_RAT
29	70	7.2	891	1	MAY1_SCHCO
30	70	7.2	972	1	SVL_MYCLE
31	70	7.2	2182	1	CABI_RAT
32	70	7.2	2261	1	ABCI_HUMAN
33	69.5	7.2	200	1	RR4_GUTH

34	69.5	7.2	356	1	ENV_FRSPB
35	69.5	7.2	998	1	EPB3_HUMAN
36	69.5	7.2	1436	1	MC11_BOVIN
37	69	7.1	180	1	DAN_HUMAN
38	69	7.1	248	1	HUTC_PSEPU
39	69	7.1	355	1	KLFE2_HUMAN
40	69	7.1	543	1	PKAA_STRCO
41	69	7.1	837	1	MUCL_RAT
42	69	7.1	1093	1	SM5B_MOUSE
43	69	7.1	1746	1	TENA_PIG
44	68.5	7.1	1056	1	MUC5_HUMAN
45	68	7.0	138	1	MUC5_CANFA

## ALIGNMENTS

RESULT 1	LSHB_MEIGA	STANDARD.	PRT:	159 AA.
ID	LSHB_MEIGA			
AC	P45646;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	LUTROPIN BETA CHAIN PRECURSOR (LUTEINIZING HORMONE BETA SUBUNIT) (LSH-BETA) (LSH-B) (LH-B).			
GN	LHB.			
OS	Meleagris gallinavo (Common turkey).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Archosauria; Aves; Neognathae; Galliformes; Meleagrididae; Meleagris.			
OX	NCBI_Taxid=9103;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Pituitary;			
RX	MEDLINE=95290073; PubMed=7772235;			
RA	You S., Foster L.K., Sliby J.L., el Halawani M.E., Foster D.N.;			
RT	"Sequence analysis of the turkey LH beta subunit and its regulation by gonadotropin-releasing hormone and prolactin in cultured pituitary cells.";			
RT	J. Mol. Endocrinol. 14:117-129(1995).			
RL	FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.			
CC	-1- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA CHAIN WHICH CONTERS BIOLOGICAL SPECIFICITY TO THYROTROPIN, LUTROPIN, FOLLITROPIN AND GONADOTROPIN.			
CC	-1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN FAMILY.			
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).			
CC	EMBL: L35519; AAA74125.1; AL1_INT1.			
DR	HSSP: P01233; 1HRP.			
DR	InterPro: IPR000359; -			
DR	InterPro: IPR001545; -			
DR	InterPro: IPR002400; -			
DR	Pfam: PF00007; Cys_knot, 1.			
DR	PRINTS: PR00438; GFCYSKNOT.			
DR	PROSITE: PS00261; GLYCO_HORMONE_BETA_1; 1.			
DR	PROSITE: PS00689; GLYCO_HORMONE_BETA_2; 1.			
KW	Hormone; Signal; Glycoprotein.			
FT	SIGNAL	1	39	POTENTIAL.
FT	CHAIN	1	159	LUTROPIN BETA CHAIN.
FT	DISULFID	48	96	BY SIMILARITY.
FT	DISULFID	62	111	BY SIMILARITY.
FT	DISULFID	65	149	BY SIMILARITY.
FT	DISULFID	73	127	BY SIMILARITY.
FT	DISULFID	77	129	BY SIMILARITY.

FT DISULFID 132 139 BY SIMILARITY  
 FT CARBOHYD 52 52 N-LINKED (GLCNAC... ) (POTENTIAL).  
 SQ SEQUENCE 159 AA; 16285 MW; 52850C8C879653C6 CRC64;

Query Match 10.3%; Score 99.5; DB 1; Length 159;  
 Best Local Similarity 27.4%; Pred. No. 0.008;  
 Matches 34; Conservative 11; Mismatches 58; Indels 21; Gaps 5;

QY 31 GGRPHHPEFTKDASEYSCRELHFTTRYVDGPCRSKAPYTELVCSGCGPARLLPNAIGR 90  
 DB 43 GGRPP-----CRPINTVAVEKEDECPQCMATTTACGGYCTREPV---EFVYR 85  
 QY 91 KWMRPSG-PDFRCIPDRYRAQVOLL-CPGGAAPRAKRVLVASCKCKRLTRFHNOSELK 149  
 DB 86 SPLGRP--PQSSCTYGALRYERMALMGCPIGSDPRV-LLPVALSCRCARCPATSDCTVQ 142  
 QY 150 DFGP 153  
 DB 143 GLGP 146

RESULT 2  
 ID LSHB\_CONJA STANDARD; PRT; 166 AA.  
 AC P45657;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE LUTROPIN BETA CHAIN PRECURSOR (LUTEINIZING HORMONE BETA SUBUNIT) (LSH-BETA) (LSH-B) (LH-B).  
 GN LHB.  
 OS Coturnix coturnix japonica (Japanese quail).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Coturnix;  
 CX NCBI\_TaxID=93934;  
 RN [1]  
 RA MEDLINE=94252550; PubMed=7515015;  
 RX Ando H., Ishii S.;  
 RT "Molecular cloning of complementary deoxyribonucleic acids for the pituitary glycoprotein hormone alpha-subunit and luteinizing hormone beta-subunit precursor molecules of Japanese quail (Coturnix coturnix japonica)."  
 RL Gen. Comp. Endocrinol. 93:357-368(1994).  
 CC -1- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.  
 CC -1- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN, LUTROPIN, FOLLITROPIN AND GONADOTROPIN.  
 CC -1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 DR EMBL: S70834; AAB30867.1; -  
 DR HSSP: P01233; IHRP.  
 DR InterPro: IPR000359; -  
 DR InterPro: IPR001545; -  
 DR Pfam: PFO0007; Cys\_knot; 1.  
 DR PRINTS: PR00438; GFCYSKNOT.  
 DR PROSITE: PS00261; GLYCO\_HORMONE\_BETA\_1; 1.  
 DR PROSITE: PS00689; GLYCO\_HORMONE\_BETA\_2; 1.  
 KW Hormone; Signal; 1  
 FT SIGNAL 1 21 POTENTIAL.

FT CHAIN 22 166 LUTROPIN BETA CHAIN.  
 FT CARBOHYD 60 60 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT DISULFID 56 104 BY SIMILARITY.  
 FT DISULFID 70 119 BY SIMILARITY.  
 FT DISULFID 73 157 BY SIMILARITY.  
 FT DISULFID 81 135 BY SIMILARITY.  
 FT DISULFID 85 137 BY SIMILARITY.  
 FT DISULFID 140 147 BY SIMILARITY.  
 SQ SEQUENCE 166 AA; 17030 MW; 6BF293BEC3C5CFAC7 CRC64;

Query Match 8.6%; Score 83.5; DB 1; Length 166;  
 Best Local Similarity 24.8%; Pred. No. 0.29;  
 Matches 31; Conservative 13; Mismatches 56; Indels 25; Gaps 5;

QY 32 GGRPHHPEFTKDASEYSCRELHFTTRYVDGPCRSKAPYTELVCSGCGPARLLPNAIGR 91  
 DB 52 GRPP-----CRPINTVAVEKEDECPQCMATTTACGGYCTREPV----- 91  
 QY 92 KWMRPSG--PDFRCIPDRYRAQVOLL-CPGGAAPRAKRVLVASCKCKRLTRFHNOSEL 148  
 DB 92 YRSLGPPPPQSSCTYGALRYERMALMGCPIGSDPRV-LLPVALSCRCARCPATSDCTV 149  
 QY 149 KDFGP 153  
 DB 150 QGLGP 154

RESULT 3  
 ID LSHB\_TRIYU STANDARD; PRT; 141 AA.  
 AC Q46482;  
 DT 01-OCT-2000 (Rel. 40, Created)  
 DT 01-OCT-2000 (Rel. 40, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE LUTROPIN BETA CHAIN PRECURSOR (LUTEINIZING HORMONE BETA SUBUNIT) (LSH-BETA) (LSH-B) (LH-B).  
 GN LHB.  
 OS Trichosurus vulpecula (Brush-tailed possum).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Metatheria; Diprotodontia; Phalangeridae; Trichosurus.  
 CX NCBI\_TaxID=9337;  
 RN [1]  
 RA SEQUENCE FROM N.A.  
 RC TISSUE-Pituitary;  
 RX MEDLINE=98345424; PubMed=9680384;  
 RX Harrison G.A., Deane E.M., Cooper D.W.;  
 RT "cDNA cloning of luteinizing hormone subunits from brushtail possum and red kangaroo."  
 RL Mamm. Genome 9:638-642(1998).  
 RN [2]  
 RA SEQUENCE FROM N.A.  
 RL Lawrence S.B., McNatty K.P., Fidler A.E.;  
 RL Submitted (SEP-1998) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.  
 CC -1- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN, LUTROPIN, FOLLITROPIN AND GONADOTROPIN.  
 CC -1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 DR EMBL: AF017448; AAC96019.1; -  
 DR EMBL: AF090388; AAC63526.1; -  
 DR HSSP: P01233; IHRP.



```

RESULT 2
Query Match          10.5%; Score 102; DB 13; Length 272;
Best Local Similarity 24.3%; Pred. No. 0.009;
Matches 26; Conservative 18; Mismatches 41; Indels 22; Gaps 4;

ID Q9PWB0 PRELIMINARY; PRT; 272 AA.
AC Q9PWB0;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE CERBERUS HOMOLOG.
GN CER.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99439862; PubMed=10508582;
RA Zhu L., Marvin M.D., Gardiner A., Lassar A.B., Mercola M., Stern C.D.,
  Levin M.;
  "Cerberus regulates left-right asymmetry of the embryonic head and
  heart.";
  Curr. Biol. 9:931-938(1999).
  EMBL: AF139721; AAD51610.1; -.
  InterPro: IPR000359; -.
  DR PROSITE: PS01225; CTCK_2; 1.
  SMART: SM00041; CT; 1.
  SQ SEQUENCE 272 AA; 31224 MW; 36E4C9F719711BCA CRC64;

Query Match          10.5%; Score 102; DB 13; Length 272;
Best Local Similarity 24.3%; Pred. No. 0.009;
Matches 26; Conservative 18; Mismatches 41; Indels 22; Gaps 4;

ID Q9PWB0 PRELIMINARY; PRT; 272 AA.
AC Q9PWB0;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE CERBERUS HOMOLOG.
GN CER.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99439862; PubMed=10508582;
RA Zhu L., Marvin M.D., Gardiner A., Lassar A.B., Mercola M., Stern C.D.,
  Levin M.;
  "Cerberus regulates left-right asymmetry of the embryonic head and
  heart.";
  Curr. Biol. 9:931-938(1999).
  EMBL: AF139721; AAD51610.1; -.
  InterPro: IPR000359; -.
  DR PROSITE: PS01225; CTCK_2; 1.
  SMART: SM00041; CT; 1.
  SQ SEQUENCE 272 AA; 31201 MW; 00D9729715771BC2 CRC64;

```

```

Query Match          10.3%; Score 100; DB 13; Length 272;
Best Local Similarity 24.3%; Pred. No. 0.014;
Matches 26; Conservative 17; Mismatches 42; Indels 22; Gaps 4;

ID Q9Y557 PRELIMINARY; PRT; 712 AA.
AC Q9Y557;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)
DE Djl163l1.3 (NOVEL PROTEIN SIMILAR TO MOUSE B99) (FRAGMENT).
GN Djl163l1.3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lloyd D.;
  Submitted (JAN-1999) to the EMBL/Genbank/DBJ databases.
  DR EMBL: AL031588; CAB38415.1; -.
  FT NON-TER
  SQ SEQUENCE 712 AA; 75685 MW; ECBD2133EF72F32 CRC64;

Query Match          9.6%; Score 93; DB 4; Length 712;
Best Local Similarity 22.4%; Pred. No. 0.2;
Matches 34; Conservative 21; Mismatches 37; Indels 60; Gaps 6;

ID Q9Y23 PRELIMINARY; PRT; 720 AA.
AC Q9Y23;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE B99.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Monte M., Collavin L., Lazarevic D., Utrera R., Schneider C.;
  "Cloning, chromosome mapping and functional characterization of a
  human homolog gene of murine B99.";
  Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
  DR EMBL: AF223408; AAF31459.1; -.
  SQ SEQUENCE 720 AA; 76628 MW; D4E852CE4BE85BBF CRC64;

```

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2001, 15:48:43 ; Search time 80.28 Seconds  
(without alignments)  
132,908 Million cell updates/sec

Title: us-09-668-021-16

Perfect score: 967  
Sequence: 1 MDATETPELGEFPEPEPEL.....RPTGKRLPRARCTKASRA 176

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

riched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq\_0601.\*  
1: /SID8/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID8/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID8/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID8/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID8/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID8/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID8/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID8/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID8/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID8/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID8/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID8/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID8/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID8/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID8/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID8/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID8/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID8/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID8/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID8/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID8/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID8/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	967	100.0	176	21	AAV96434
2	920.5	95.2	213	21	AAV96431
3	901.5	93.2	213	21	AAV96429
4	901.5	93.2	213	21	AAV96429
5	901.5	93.2	213	21	AAV96429
6	892.5	92.3	213	21	AAV96430
7	892.5	92.3	213	21	AAV96436
8	861.5	89.1	213	21	AAV96433
9	857.5	88.7	211	21	AAV96432
10	808.5	83.6	367	21	AAV96432
11	336	34.7	206	21	AAV96432

12	336	34.7	206	21	AAV96431	Murine skin cell p
13	336	34.7	206	22	AAV96431	Skin cell protein,
14	336	34.7	206	22	AAV96431	Skin cell protein,
15	336	34.7	206	18	AAV96431	Human small CCN-11
16	336	34.4	206	19	AAV96431	Human small CCN-11
17	336	34.4	206	20	AAV96431	Homo sapiens fetal
18	336	34.4	206	21	AAV96431	Human adult retina
19	108.5	11.2	116	18	AAV96431	Secreted protein A
20	108.5	11.2	116	18	AAV96431	Human secreted pro
21	103.5	10.7	102	20	AAV96431	Amino acid sequenc
22	103.5	10.7	102	21	AAV96431	Human fetal kidney
23	96	9.9	50	20	AAV96431	Human 5' EST seque
24	93	9.6	712	21	AAV96431	Human ORF ORF2911
25	87	9.0	168	21	AAV96431	Amino acid sequenc
26	86	8.9	168	21	AAV96431	Murine cerebral ne
27	85.5	8.8	320	20	AAV96431	Mycobacterium spec
28	85	8.8	184	19	AAV96431	DAN and b57 protei
29	83.5	8.6	1428	21	AAV96431	Caspase 8-interact
30	82.5	8.5	145	12	AAV96431	hCG/ELH chimera, E
31	80.5	8.3	212	20	AAV96431	Human CG beta subu
32	79.5	8.2	1480	13	AAV96431	hCG/ELH chimera, E
33	79	8.2	212	20	AAV96431	Drosophila SLIT pr
34	79	8.2	246	21	AAV96431	hCG/hFSH beta subu
35	79	8.2	470	20	AAV96431	Herpes simplex vir
36	79	8.2	470	21	AAV96431	Human repressor Ki
37	79	8.2	470	22	AAV96431	Human RFLF protein
38	79	8.2	479	20	AAV96431	Human RFLF protein
39	79	8.2	479	21	AAV96431	Human repressor Ki
40	79	8.2	685	20	AAV96431	Human RFLF protein
41	79	8.2	685	20	AAV96431	Disease associated
42	78.5	8.1	204	20	AAV96431	Human CG beta subu
43	77.5	8.0	145	12	AAV96431	hCG/ELH chimera, E
44	77.5	8.0	145	18	AAV96431	Chorionic gonadotr
45	77.5	8.0	184	20	AAV96431	Human DM protein

#### ALIGNMENTS

RESULT 1	AAV96434	standard: Protein; 176 AA.
ID	AAV96434	
XX	AAV96434	
AC	AAV96434	
XX	12-SEP-2000 (first entry)	
DT	XX	
DE	XX	Bovine TGF-beta binding protein (BEER).
XX	XX	osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
KW	BEER; gene therapy; antisense therapy; fracture; bone mineralization.	
XX	OS	Bos taurus.
XX	PN	WO200032773-A1.
XX	PN	08-JUN-2000.
XX	PD	24-NOV-1999; 99WO-US27990.
XX	PF	24-NOV-1999; 99WO-US27990.
XX	PR	27-NOV-1998; 98US-0110283.
XX	XX	(DARW-) DARWIN DISCOVERY LTD.
PA	XX	Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;
XX	PI	Van Ness J, Winkler DG;
PI	XX	WPI: 2000-412321/35.
DR	XX	N-PSDB; AAA29060.
DR	XX	Nucleic acids (I) encoding a transforming growth factor beta binding
PT	XX	protein, useful for identifying agents for treating osteopenia,
PT	XX	osteoporosis and fractures

XX Claim 7, Page 127, 162pp; English.  
 PS  
 CC This shows a bovine transforming growth factor-beta (TGF-beta)  
 CC binding protein designated bBER. The cDNA and protein may be used for  
 CC prevention, treatment and diagnosis of diseases associated with  
 CC inappropriate BBR expression. For example, they may be used to treat  
 CC disorders associated with decreased TGF-beta BP expression. The cDNA or  
 CC vectors may be administered to treat diseases by rectifying mutations or  
 CC deletions in a patient's genome that affect the activity of BBR by  
 CC expressing inactive proteins or to supplement the patients own production  
 CC of BBR polypeptides. The nucleic acids may be used for recombinant  
 CC production of BBR, gene therapy, antisense therapy, as probes for  
 CC diagnostic assays and for functional studies. BBR may be used to raise  
 CC antibodies and for identification of BBR modulators. BBR antagonists  
 CC may be used to increase bone mineral content for the treatment of  
 CC disorders such as osteopenia, osteoporosis, fractures and other disorders  
 CC associated with low mineral content.  
 XX  
 SQ Sequence 176 AA;

Query Match 100.0%; Score 967; DB 21; Length 176;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-87;  
 Matches 176; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NDATETIIPELGEYPEPLPELNKTKMRAENGRRPHHPTETDASRYSCRELIHFTRYTD 60  
 DB 1 ndatetipelgeypeplpeelnktkmraengrrphhptetkdasyscrelihftrytd 60  
 OY 61 GPCRSKAPVTELVCSGCGPARLLPNAIGRGKWMRSPGDFPCIDPRYRAQRYVQLLCPG 120  
 DB 61 gpcrsakpvtelvcsgcgparllpnaigrkwmpsgdfcicpdyraqrvqllcp99 120  
 OY 121 AABRAKRVLVASCKCKRLTRFHNOSELKDFGEARPOGTGKRLPRRAGTAAASA 176  
 DB 121 aabrakrvlvascckckrltrfhngselkdfgeaarpgtgkrlprargtaasra 176

RESULT 2  
 ID AAY96431 standard; Protein; 213 AA.  
 AC AAY96431;  
 DT 12-SEP-2000 (first entry)

DE Vervet TGF-beta binding protein (bBER).  
 XX

osteopathic; transforming growth factor-beta; TGF-beta; binding protein;  
 bBER; gene therapy; antisense therapy; fracture; bone mineralization.

OS Cercopithecus pygerythrus.

XX WO200032773-A1.  
 XX 08-JUN-2000.  
 XX 24-NOV-1999; 99WO-US27990.  
 XX 27-NOV-1998; 98US-0110283.  
 XX (DARW-) DARWIN DISCOVERY LTD.  
 XX PA  
 XX PI Brunckow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;  
 XX PI Van Ness J, Winkler DG;  
 XX DR WPI; 2000-412321/35.  
 XX DR N-PSDB; AAA29057.

XX Nucleic acids (I) encoding a transforming growth factor beta binding  
 PT protein, useful for identifying agents for treating osteopenia,  
 PT osteoporosis and fractures

XX Claim 4, Page 122-123, 162pp; English.  
 PS  
 CC This shows a vervet transforming growth factor-beta (TGF-beta)  
 CC binding protein designated vBER. The cDNA and protein may be used for  
 CC prevention, treatment and diagnosis of diseases associated with  
 CC inappropriate BBR expression. For example, they may be used to treat  
 CC disorders associated with decreased TGF-beta BP expression. The cDNA or  
 CC vectors may be administered to treat diseases by rectifying mutations or  
 CC deletions in a patient's genome that affect the activity of BBR by  
 CC expressing inactive proteins or to supplement the patients own production  
 CC of BBR polypeptides. The nucleic acids may be used for recombinant  
 CC production of BBR, gene therapy, antisense therapy, as probes for  
 CC diagnostic assays and for functional studies. BBR may be used to raise  
 CC antibodies and for identification of BBR modulators. BBR antagonists  
 CC may be used to increase bone mineral content for the treatment of  
 CC disorders such as osteopenia, osteoporosis, fractures and other disorders  
 CC associated with low mineral content.  
 XX  
 SQ Sequence 213 AA;

Query Match 95.2%; Score 920.5; DB 21; Length 213;  
 Best Local Similarity 95.5%; Pred. No. 7.2e-83;  
 Matches 169; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

OY 1 NDATETIIPELGEYPEPLPELNKTKMRAENGRRPHHPTETDASRYSCRELIHFTRYTD 59  
 DB 31 ndatetipelgeypeplpeelnktkmraengrrphhptetkdasyscrelihftrytd 90  
 OY 60 DGPCRSKAPVTELVCSGCGPARLLPNAIGRGKWMRSPGDFPCIDPRYRAQRYVQLLCPG 119  
 DB 91 dgpcrsakpvtelvcsgcgparllpnaigrkwmpsgdfcicpdyraqrvqllcp9 150  
 OY 120 GAAPRAKRVLVASCKCKRLTRFHNOSELKDFGEARPOGTGKRLPRRAGTAAASA 176  
 DB 151 gaaprakrvlvascckckrltrfhngselkdfgeaarpgtgkrlprargtaasra 207

RESULT 3  
 ID AAB26106 standard; Protein; 213 AA.  
 AC AAB26106;  
 DT 15-JAN-2001 (first entry)

DE Human DAN/Cerberus-related protein 6 (hDCR6) (exons 1 and 4).

XX Human; DNA/Cerberus-related protein 6; hDCR6; morphogenic protein;  
 KW antagonist; BMP; cell growth; cell differentiation; bone formation;  
 XX gene therapy.  
 XX

OS Homo sapiens.

XX WO200055193-A2.  
 XX 21-SEP-2000.  
 XX 02-MAR-2000; 2000WO-US05537.  
 XX 12-MAR-1999; 99US-0124118.  
 XX (REGG-) REGENERON PHARM INC.  
 XX PA  
 XX PI Economides AN;  
 XX PI  
 XX DR WPI; 2000-638179/61.  
 XX DR N-PSDB; AAA94051.

XX Novel isolated, human DNA/Cerberus related protein 6 which include  
 PT natural homologue, and polypeptides comprising DCR6 domain and nucleic  
 PT acids encoding the proteins which are useful as probes and primers







GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 8, 2001, 19:18:35 ; Search time 5549.44 Seconds  
(without alignments)  
3919.497 Million cell updates/sec

Title: US-09-668-021-1

Perfect score: 2301  
Sequence: 1 agagcctgtctacttgaag.....caatgaatcatgaccgaaag 2301

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Indexed: 1022815 segs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*  
10: gb\_est10:\*  
11: gb\_est11:\*  
12: gb\_est12:\*  
13: gb\_est13:\*  
14: gb\_est14:\*  
15: gb\_est15:\*  
16: gb\_est16:\*  
17: gb\_est17:\*  
18: gb\_est18:\*  
19: gb\_est19:\*  
20: gb\_est20:\*  
21: gb\_est21:\*  
22: gb\_est22:\*  
23: gb\_est23:\*  
24: gb\_est24:\*  
25: gb\_est25:\*  
26: gb\_est26:\*  
27: gb\_est27:\*  
28: gb\_est28:\*  
29: gb\_est29:\*  
30: gb\_est30:\*  
31: gb\_est31:\*  
32: gb\_est32:\*  
33: gb\_est33:\*  
34: gb\_est34:\*  
35: gb\_est35:\*  
36: gb\_est36:\*  
37: gb\_est37:\*  
38: gb\_est38:\*  
39: gb\_est39:\*  
40: gb\_est40:\*  
41: gb\_est41:\*  
42: gb\_est42:\*  
43: gb\_est43:\*

44: em\_esthum10:\*  
45: em\_esthum11:\*  
46: em\_esthum12:\*  
47: em\_esthum13:\*  
48: em\_esthum14:\*  
49: em\_esthum15:\*  
50: em\_esthum16:\*  
51: em\_esthum17:\*  
52: em\_esthum18:\*  
53: em\_esthum19:\*  
54: em\_esthum20:\*  
55: em\_esthum21:\*  
56: em\_esthum22:\*  
57: em\_esthum23:\*  
58: em\_esthum24:\*  
59: em\_esthum25:\*  
60: em\_esthum26:\*  
61: em\_esthum27:\*  
62: em\_esthum28:\*  
63: em\_esthum29:\*  
64: em\_esthum30:\*  
65: em\_esthum31:\*  
66: em\_esthum32:\*  
67: em\_esthum33:\*  
68: em\_esthum34:\*  
69: em\_esthum35:\*  
70: em\_esthum36:\*  
71: em\_esthum37:\*  
72: em\_esthum38:\*  
73: em\_esthum39:\*  
74: em\_esthum40:\*  
75: em\_esthum41:\*  
76: em\_esthum42:\*  
77: em\_esthum43:\*  
78: em\_esthum44:\*  
79: em\_esthum45:\*  
80: em\_esthum46:\*  
81: em\_esthum47:\*  
82: em\_esthum48:\*  
83: em\_esthum49:\*  
84: em\_esthum50:\*  
85: em\_esthum51:\*  
86: em\_esthum52:\*  
87: em\_esthum53:\*  
88: em\_esthum54:\*  
89: em\_esthum55:\*  
90: em\_esthum56:\*  
91: em\_esthum57:\*  
92: em\_esthum58:\*  
93: em\_esthum59:\*  
94: em\_esthum60:\*  
95: em\_esthum61:\*  
96: em\_esthum62:\*  
97: em\_esthum63:\*  
98: em\_esthum64:\*  
99: em\_esthum65:\*  
100: em\_esthum66:\*  
101: em\_esthum67:\*  
102: em\_esthum68:\*  
103: em\_esthum69:\*  
104: em\_esthum70:\*  
105: em\_esthum71:\*  
106: em\_esthum72:\*  
107: em\_esthum73:\*  
108: em\_esthum74:\*  
109: em\_esthum75:\*  
110: em\_esthum76:\*  
111: em\_esthum77:\*  
112: em\_esthum78:\*  
113: em\_esthum79:\*  
114: em\_esthum80:\*  
115: em\_esthum81:\*  
116: em\_esthum82:\*

117: gb\_est48:\*  
118: gb\_est49:\*  
119: gb\_est50:\*  
120: gb\_est51:\*  
121: gb\_est52:\*  
122: gb\_est53:\*  
123: gb\_est54:\*  
124: gb\_est55:\*  
125: gb\_est56:\*  
126: gb\_est57:\*  
127: gb\_est58:\*  
128: gb\_est59:\*  
129: gb\_est60:\*  
130: gb\_est61:\*  
131: gb\_est62:\*  
132: gb\_est63:\*  
133: gb\_est64:\*  
134: gb\_est65:\*  
135: gb\_est66:\*  
136: gb\_est67:\*  
137: gb\_est68:\*  
138: gb\_est69:\*  
139: gb\_est70:\*  
140: gb\_est71:\*  
141: gb\_est72:\*  
142: gb\_est73:\*  
143: gb\_est74:\*  
144: gb\_est75:\*  
145: gb\_est76:\*  
146: gb\_est77:\*  
147: gb\_est78:\*  
148: gb\_est79:\*  
149: gb\_est80:\*  
150: gb\_est81:\*  
151: gb\_est82:\*  
152: gb\_est83:\*  
153: gb\_est84:\*  
154: gb\_est85:\*  
155: gb\_est86:\*  
156: gb\_est87:\*  
157: gb\_est88:\*  
158: gb\_est89:\*  
159: gb\_est90:\*  
160: gb\_est91:\*  
161: gb\_est92:\*  
162: gb\_est93:\*  
163: gb\_est94:\*  
164: gb\_est95:\*  
165: gb\_est96:\*  
166: gb\_est97:\*  
167: gb\_est98:\*  
168: gb\_est99:\*  
169: gb\_est100:\*  
170: gb\_est101:\*  
171: gb\_est102:\*  
172: gb\_est103:\*  
173: gb\_est104:\*  
174: gb\_est105:\*  
175: gb\_est106:\*  
176: gb\_est107:\*  
177: gb\_est108:\*  
178: gb\_est109:\*  
179: gb\_est110:\*  
180: gb\_est111:\*  
181: gb\_est112:\*  
182: gb\_est113:\*  
183: gb\_est114:\*  
184: gb\_est115:\*  
185: gb\_est116:\*  
186: gb\_est117:\*  
187: gb\_est118:\*  
188: gb\_est119:\*  
189: gb\_est120:\*

190: gb\_est110:\*  
191: gb\_est111:\*  
192: gb\_est112:\*  
193: gb\_est113:\*  
194: gb\_est114:\*  
195: gb\_est115:\*  
196: gb\_est116:\*  
197: gb\_est117:\*  
198: gb\_est118:\*  
199: gb\_est119:\*  
200: gb\_est120:\*  
201: gb\_est121:\*  
202: gb\_est122:\*  
203: gb\_est123:\*  
204: gb\_est124:\*  
205: gb\_est125:\*  
206: gb\_est126:\*  
207: gb\_est127:\*  
208: gb\_est128:\*  
209: gb\_est129:\*  
210: gb\_est130:\*  
211: gb\_est131:\*  
212: gb\_est132:\*  
213: gb\_est133:\*  
214: gb\_est134:\*  
215: gb\_est135:\*  
216: gb\_est136:\*  
217: gb\_est137:\*  
218: gb\_est138:\*  
219: gb\_est139:\*  
220: gb\_est140:\*  
221: gb\_est141:\*  
222: gb\_est142:\*  
223: gb\_est143:\*  
224: gb\_est144:\*  
225: gb\_est145:\*  
226: gb\_est146:\*  
227: gb\_est147:\*  
228: gb\_est148:\*  
229: gb\_est149:\*  
230: gb\_est150:\*  
231: gb\_est151:\*  
232: gb\_est152:\*  
233: gb\_est153:\*  
234: gb\_est154:\*  
235: gb\_est155:\*  
236: gb\_est156:\*  
237: gb\_est157:\*  
238: gb\_est158:\*  
239: gb\_est159:\*  
240: gb\_est160:\*  
241: gb\_est161:\*  
242: gb\_est162:\*  
243: gb\_est163:\*  
244: gb\_est164:\*  
245: gb\_est165:\*  
246: gb\_est166:\*  
247: gb\_est167:\*  
248: gb\_est168:\*  
249: gb\_est169:\*  
250: gb\_est170:\*  
251: gb\_est171:\*  
252: gb\_est172:\*  
253: gb\_est173:\*  
254: gb\_est174:\*  
255: gb\_est175:\*  
256: gb\_est176:\*  
257: gb\_est177:\*  
258: gb\_est178:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 8, 2001, 23:51:52 ; Search time 438.73 Seconds  
(without alignments)  
918.818 Million cell updates/sec

Title: us-09-668-021-9

Perfect score: 642  
Sequence: 1 atcgagctccacgtgcctt.....agctgagagagcctactag 642

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searches: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

N\_Geneseq\_0601.\*  
1: /SIDSI/gcgdata/geneseq/geneseqn/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseqn/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseqn/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseqn/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseqn/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseqn/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseqn/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseqn/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseqn/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseqn/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseqn/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseqn/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseqn/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseqn/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseqn/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseqn/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseqn/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseqn/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseqn/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseqn/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseqn/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseqn/NA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	642	100.0	642	21	AAA29057
2	613.2	95.5	642	21	AAA29051
3	613.2	95.5	2301	21	AAA29055
4	611.6	95.3	2329	22	AAA29053
5	611.6	95.3	2301	21	AAA29061
6	611.6	95.3	2301	21	AAA29062
7	610	95.0	2301	21	AAA29056
8	530	82.6	674	21	AAA29059
9	503.2	78.4	638	21	AAA29058
10	461.4	71.9	532	21	AAA29060
11	413.4	64.4	5680	21	AAA29049

12	413.4	64.4	9301	21	AAA29064	Human TGF-beta bin
13	399.6	62.2	1104	21	AAA29050	Human DAN/Cerberus
14	361	56.2	35828	21	AAA29063	Murine TGF-beta b1
15	142	22.1	198	20	AA240842	Human secreted pro
16	141.6	22.1	196	21	AAC05741	Human secreted pro
17	83.4	13.0	954	22	AAC99782	Skin cell cDNA, SE
18	83.4	13.0	962	21	AA261645	CDNA encoding murt
19	83.4	13.0	962	21	AA261736	CDNA encoding murt
20	83.4	13.0	962	22	AAC99578	CDNA encoding murt
21	83.4	13.0	962	22	AAC99669	CDNA encoding murt
22	71	11.1	900	18	AAT47661	Human small CCN-11
23	71	11.1	900	19	AAV37735	CDNA sequence of h
24	71	11.1	1692	21	AAC79467	Human adult retina
25	71	11.1	1710	20	AA284238	Human adult retina
26	71	11.1	1719	21	AA284238	Human adult retina
27	71	11.1	1774	20	AA284238	Human adult retina
28	57.8	9.0	412	22	AA284238	Human adult retina
29	54	8.4	37856	21	AA284238	Human adult retina
30	52.2	8.1	2943	17	AAT16480	Novel human polyu
31	48.2	7.5	114955	20	AA284238	S. cellulosum DNA
32	47.6	7.4	114955	20	AA284238	Human adenosine A1
33	47.2	7.4	71889	21	AA284238	Sorngium cellulos
34	45.4	7.1	23666	12	AAQ10190	Cephalosporin anti
35	44.2	6.9	835	18	AAT77280	Lymantria dispar n
36	44.2	6.9	2791	16	AAO99776	Phosphorylase biosy
37	44.2	6.9	8867	21	AA284238	Human Lhx3 genomic
38	43.6	6.8	5692	22	AA284238	Streptomyces sp. C
39	43.4	6.8	1125	22	AA284238	Signal transductio
40	43	6.7	4020	18	AAT91361	Orf virus genomic
41	42.6	6.6	4020	18	AAT91361	Orf virus genomic
42	42.6	6.6	5617	19	AAV57163	Partial human Notc
43	42.6	6.6	8091	19	AAV57001	Human Notch3 cDNA
44	42.6	6.6	43280	18	AAT80413	Tylosine synthase
45	42	6.5	117213	19	AA284238	HSV-2 strain SB5 C

#### ALIGNMENTS

RESULT	1	
ID	AAA29057	standard; cDNA: 642 BP.
AC	AAA29057;	
DT	12-SEP-2000	(first entry)
DE	Vervet TGF-beta binding protein (BEER) cDNA.	
KW	osteopathic; transforming growth factor-beta; TGF-beta; binding protein;	
KM	BEER; gene therapy; antisense therapy; fracture; bone mineralization; ss.	
OS	Cercopithecus pygerythrus.	
PH	Key	Location/Qualifiers
FT	CDS	1..642
FT		/*tag= a
XX		/product= TGF-beta_binding_protein
PN	WO200032773-A1.	
PD	08-JUN-2000.	
XX		
XX	24-NOV-1999;	99WO-US27990.
PF		
XX		
PR	27-NOV-1998;	98US-0110283.
XX		
PA	(DARW-) DARWIN DISCOVERY LTD.	
XX		
PI	Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;	
XX	Van Ness J, Winkler DG;	
DR	WPI; 2000-412321/35.	



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 8, 2001, 19:20:30 ; Search time 9526.57 Seconds  
(without alignments)  
3736.003 Million cell updates/sec

Title: US-09-668-021-1

Perfect score: 2301  
Sequence: 1 agagcctgtgcactcgaag.....caatgaatcatcaccgaag 2301

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Checked: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: \*  
1: gb\_ba1: \*  
2: gb\_ba2: \*  
3: gb\_ba3: \*  
4: gb\_in1: \*  
5: gb\_in2: \*  
6: gb\_in3: \*  
7: gb\_om: \*  
8: gb\_ov: \*  
9: gb\_pat1: \*  
10: gb\_pat2: \*  
11: gb\_ph: \*  
12: gb\_pl1: \*  
13: gb\_pl2: \*  
14: gb\_pl3: \*  
15: gb\_pl4: \*  
16: em\_ba1: \*  
17: em\_ba2: \*  
18: em\_fun: \*  
19: em\_htgo\_hum: \*  
20: em\_htgo\_inv: \*  
21: em\_htgo\_inv: \*  
22: em\_htg\_hum1: \*  
23: em\_htg\_hum2: \*  
24: em\_htg\_hum3: \*  
25: em\_htg\_hum4: \*  
26: em\_htg\_hum5: \*  
27: em\_htg\_hum6: \*  
28: em\_htg\_hum7: \*  
29: em\_htg\_hum8: \*  
30: em\_htg\_inv1: \*  
31: em\_htg\_inv2: \*  
32: em\_htg\_other: \*  
33: em\_htg\_rtd: \*  
34: em\_hum1: \*  
35: em\_hum2: \*  
36: em\_hum3: \*  
37: em\_hum4: \*  
38: em\_hum5: \*  
39: em\_hum6: \*  
40: em\_hum7: \*  
41: em\_in: \*  
42: em\_om: \*  
43: em\_or: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2301	100.0	2323	89	AF326739 Homo sapi
2	2280.6	99.1	2329	9	AX056687 Sequence
3	2273.2	98.8	2296	89	AF331844 Homo sapi
4	2047.8	89.0	21501	89	AF326736 Homo sapi
5	2047.8	89.0	94752	85	AC003098 Homo sapi
6	2035.8	88.5	151780	72	AC055813 Homo sapi
7	1066.4	46.3	177744	75	AC073954 Homo sapi
8	613.2	26.6	642	89	AF326742 Cercopithec





QY 781 attcatgttaaalgtcctgcaaccgagggagctgagacccctccagagcccttgaga 840  
 DB 781 ATTTCATTGTAATAATGCGTCAACCCAGGGAGGGGGCTGAGACCTTCAGGCGCTGAGGA 840  
 QY 841 atcccgagggcgccgcaagagcccccctcaagcccgcaagctgaggggtcccaagggcgagg 900  
 DB 841 ATCCCGGCGCGCCGCAAGGGCCCCCTCAGCCCGCAGCTGAGGGGTCCACAGGGGGAGG 900  
 QY 901 gagggaaattgagatcagacagactgagccagcccgctctgaggcgccactaact 960  
 DB 901 GAGGGAATGAGATGACAGACACTGAGCCAGCCAGCCCGCTCTGAGGCGCCCTACT 960  
 QY 961 ctgctgtgtcccaacttcagagagagaaatggaagaactttccagccgctggggttta 1020  
 DB 961 TTGCTGTGTCCTTCCACTTCAGAGAGGAGGAAATGGAAGCATTTTCACCGCGCTGGGTTTA 1020  
 QY 1021 agggagcggtgtgaggagtgaggaagtcagggagctggttaagaaggttgatlaattcc 1080  
 DB 1021 AGGGAGCGGCTGTGGAGCTGGGAAAGTCAGAGGACTGGTTAAAGATTGAGTAAAGATTCC 1080  
 QY 1081 cccttgacactcgcgtgcccacacagaaagcctgagggctgcccagagcacaagactggagg 1140  
 DB 1081 CCTTGCACGCTCGCTGCGCCATCAGAAAGCCTGAGGGGTGCCAGAGCAGCAAGACTGGGGG 1140  
 QY 1141 caactgtagatgtgtgttctagctcgtgctctgccaacttctgtgtgtaacttgaaac 1200  
 DB 1141 CAACGTGATGATGTGTGTTCAGTCGTGCTGTGCCACTAACTTGTGTGTAACCTTTGAAAC 1200  
 QY 1201 tacacaattctccttcggagaccacaaattccacttgtlaaagtgagggtlgaaggltgaa 1260  
 DB 1201 TACACAATTCCTCTTGGGAGCTCAATTTCCACTTTGTAAATGAGGGTGGAGGAGGAA 1260  
 QY 1261 tagatctcgaagagactatgtgcatatattccaaagagactcaatgacctttgaatggg 1320  
 DB 1261 TAGATCTCGAGAGACTATGTGCATATGATATTCACAGGACTCAAGTCTTTGATATGG 1320  
 QY 1321 cagaggttgag 1380  
 DB 1321 CAGAGGTGAG 1380  
 QY 1381 caaaggtacactccagaaattcagaggtgtgatgtctctcttcagagcacaagaatgaa 1440  
 DB 1381 CAAGGTACTCTCCAGAAATTCAGAGGTGTGATGTCTCTCTTCAGAGCCCAAGAGTCAAAA 1440  
 QY 1441 caaacaagaaaaaaagtaagaagctatattatgtgtgacatatttcaggtctgacaac 1500  
 DB 1441 CAAACAGAAAAAAGTAAGAGTCTATTATGTGCTACATATTATTCAGGCTGCAAAAC 1500  
 QY 1501 tccctgagaagagctatgtctctccagcctgctcccgagatgttttgcttaactccac 1560  
 DB 1501 TCCCTGAGAGAGCTATGTCTCTCCAGCCTGCTCCCGAGATGTTTGGGTACCTCCAC 1560  
 QY 1561 cccctcatctcaaaagaaataacatcatcttgagggttagaaaagagagaggttcgaggg 1620  
 DB 1561 CCCCCTCATCTCAAAAGAAATACATATCATTTGGGGTAAAGAGAGAGGGGCCAGAGGT 1620  
 QY 1621 ggtggagagagtagaatacatcccgcccaacttcccaagaagagagatccctccccc 1680  
 DB 1621 GGTGGAGAGATAGAAATACATCCCGCCCACTCCCAAGAGAGAGATCCCTCCCCCG 1680  
 QY 1681 acccatagcagatgtttaaagtcaacttcgagagagaaatgaaagtgttcaagagacagg 1740  
 DB 1681 ACCCATAGCAGATGTTTAAAGTCACTTCGAGAGAGAAATGTAAGGTTTCAAGGAGACTGG 1740  
 QY 1741 ccttgcaagcccgagggagagccatcaacaactcacaagcagagacatcccttttga 1800  
 DB 1741 CTTTGCAGGCCCCGAGGAGAGCCATCAAACTCACAAGCAGACACATCCCTTTTGAGA 1800  
 QY 1801 cagcgactctgcccacacactcaagagacacattctgcttagaaaaagagcttctaagc 1860  
 DB 1801 CAGCGACTCTTCTGCCACACTCAAGGAGACATTTCTTGCTTAGAAAAACAGCTTCTTACTGC 1860

QY 1861 tcttacatgtgatgcatatcttacactaaagaatattatgtgggggaaanaactacaagt 1920  
 DB 1861 TCTTACATGTGATGGCATATCTTACACTAAAGAAATATTATTGGGGGAAAAACATCAAAAGT 1920  
 QY 1921 gctgtacatattgttgaagaactgcagagacataatagctgcaccccaaaaatcttttga 1980  
 DB 1921 GCTGTACATATGCTGAAACACTGCAGACATATAGCTGCCACCCAAAAATCTTTTGA 1980  
 QY 1981 aatcatttccagaacaacttacttctctgtgtagttttaaattgtlaaaaaaaaagt 2040  
 DB 1981 AATCATTTCAGAACACCTCTTACTTCTGTAGTTTAAATGTTTAAAAAAAAGT 2040  
 QY 2041 tttaacagagacacatagacataatgaagacctgcagagactgctgttttttgcacatc 2100  
 DB 2041 TTTAACAGAGACACATGATATGAAGCCGACGAGACTGCTGTTTTTGTGCAATTC 2100  
 QY 2101 ttccagctgggactgttccacaagaatgaagtagtgtttttaaagatgaattacat 2160  
 DB 2101 TTCCAGCTGGGACTGTGCCAAGAAATGAAGTAGTGTTTTAAAGATTAACTTACAT 2160  
 QY 2161 attatttctcacttaagttatattatgcaaaagtttctctgtagaagatgacaatgtc 2220  
 DB 2161 ATTATTTCCTCACTTACATTAATTTATGCAAAAGTTTCTGTGAGAAATGACAAATGT 2220  
 QY 2221 aatatgcttattgaattaaacagctgtctctccagagctccagagacattgttaataag 2280  
 DB 2221 AATATGCTTATGAATTAAACAGCTGTGTTCTTCAGAGTCCAGAGACTATSTTATTAAG 2280  
 QY 2281 acaatgaatcagacaggaag 2301  
 DB 2281 ACAATGAATCATGACCGAAG 2301

RESULT 2  
 AX056687  
 LOCUS AX056687 2329 bp DNA PAT 17-JAN-2001  
 DEFINITION Sequence 19 from Patent W00075317.  
 ACCESSION AX056687  
 VERSION AX056687.1 GI:12309667  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 2329)  
 AUTHORS Bolstein,D.A., Goddard,A., Gurney,A.L., Smith,V., Watanabe,C.K. and Wood,W.I.  
 TITLE Compositions and methods for the treatment of tumor  
 JOURNAL Patent: WO 0075317-A 19 14-DEC-2000;  
 Genentech, Inc. (US)  
 FEATURES  
 source location/Qualifiers  
 1..2329  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 BASE COUNT 587 a 645 c 612 g 485 t  
 ORIGIN

Query Match 99.1%; Score 2280.6; DB 9; Length 2329;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 2283; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 14 ctggaaggtgagtgagcttctctctgtgtgtacatagcagctcccaatggccgtgtct 73  
 DB 28 CTGGAAGGTGAGTGAGCTCTCTCTGTGTGTACATGAGCTCCCACTGGCCGTGTGCT 87  
 QY 74 cgtgcctctgtgtlaacacagacacttccgtgtagtgtgaagggccaggggtgtgcagcg 133  
 DB 88 CGTGTGCTGTGTGTACACACACTTCCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 147  
 QY 134 caagatatttcacagaaatcaatcccgagctcgagaggttaccgagcctccacagga 193  
 DB 148 CAGGATATGCCAGGAATCATTCCTCCGAGCTCGAGAGTACGCCGAGCTCCACCGGA 207

[illegible]

QY	1274	agacatcttgacatgatctccaaagacacccagtcgccttttgaatggcgaaagttgagaa	13
Db	1288	AGACNTATGGCATATGATTTCCAAAGACATCCAGTGCCTTTTGAATGGCGAAGGTGAGAA	1347
QY	1334	gagagagagaagagagagaatgaaatgcagtgcatttgatctcagttgcgaagttcactcc	1392
Db	1348	GAGGAGAGAAAGAGAGAAATGAAATGCATTGCTATGTTAGTGGCCAAAGTGCACCTTTC	1407
QY	1394	agaatcagaagtttgatgcagtcctctcttcgcagcccaagatgaaaaacaaacagaaanaa	1453
Db	1408	AGAAATCAGATTGTGATGCTCTCTTCTTGACACCCAAAGATGAAAAACAAACGAAAAA	1467
QY	1454	aaaagtaagagttcatttatgctgtgacatttaagcgtgaacaaactcttgaaanaagc	1513
Db	1468	AAAAGTAAGAGCTAATTTATGCTGACATATTTACGCTGACAAACCTCTGGAAGAAAGC	1527
QY	1514	tatgctgctccagccgctgagttccccgagatgttttgctacctccacccctccatctcaa	1577
Db	1528	TATGCTGCTTCCACCTCGGCTTCCCGAGATGTTTGGTACTCCACCTCCATGCTCA	1587
QY	1574	agaaataaacatcatctctggagtgtaagaaagagagaggtccgaaggtgctggagaggata	1633
Db	1588	AGAAATATCATCTATCTTTGGGGTNGAAAAGAGAGAGGTCCGAGGGTGTGGAGGGAGA	1647
QY	1634	gaatcacalcgcgcccaacttcccaaaagacgaatccctcccccagccatagcatalg	1693
Db	1648	GAATCACAATCCGCCCCAAATTTCCAAAGAGCGCATCTCTCCCGACCCATATGCGATG	1707
QY	1694	ttttaaagtaacctccgcgaagaagatgaaagtttaagaagcaatggcctctgaagcccg	1753
Db	1708	TTTTAAAGTACCTTCCGAGAGAAAGTGAAGGTTCAAGACATCGGCTTTGACGCCCG	1767
QY	1754	agggagcagccatcacaaactcaagacgcacatcccttttgagaccggcctctgc	1813
Db	1768	AGGGAGAGCGCATCAAAACTCAAGACACGACATCCCTTTGAGACCGCCTTCTGC	1827
QY	1814	ccacacatcaacggaacatctctgcgcagaaacacagcttcttactgctcttacaatgtagt	1877
Db	1828	CCACACATCAGGACACATTTCTGCTGAGAAAACACTTCTTACTGCTTTACATGTGAT	1887
QY	1874	ggcatctcttaacactaaagaatacttatctggggaaaaactacaagctgcatactagc	1933
Db	1888	GGCATCTTACACGTAAAGAAATTTATTTGGGGAAAAACATCAAGTGTGATCATATGC	1947
QY	1934	tgaagaacatgcagagcataatagctgcaccaccaaaaactctttgaaatcatcttccaga	1993
Db	1948	TGAAGAACTGCGAGCATTAATAGTGTGCACCCCAAAATCTTTTGAAATCATTTTCAAA	2007
QY	1994	caaccccttactctctgcgtatgttcttcaatctgtctaaaaaaaaagtttcaaacagaagc	2053
Db	2008	CAACCTCTTACTTTCGTGTGATGTTTAAATTTGTTAAAAAAGTTTAAACAGAAAC	2067
QY	2054	acatgacataatgaaagcctgcgaagcagtcgctcttttttggcaatcttcccaagtgagc	2113
Db	2068	ACATGACATTAATAAGCCTGAGAGCATGGTCTGTTTTTGGCAATCTTCCACAGTGGAC	2127
QY	2114	ctgtccacaagaatgaaagtagtggttctttaaagagttaagttlaacatattatcttctca	2177
Db	2128	TTGTCCACAAGAAATGAAGTGTGTTTTTAAAGAGTTAAGTTACATATTTATTTCTCA	2187
QY	2174	cttaagttattatactaaagtttcttctgtgtagaagatgaaacatgtaatactgctctat	2233
Db	2188	CTTAAGTTATTTATCCAAAGTTTTTCTGTATAGAAATGACAAATGTTAATATGGCTTAT	2247
QY	2234	gaatlaacagttctgtcttccagagttccagagacatgttlaataaagacaatgaaatcalg	2293
Db	2248	GAATTAACAGTCTGTGTTCCAGAGTCCAGAGACATTTGTAATTAAGAACATGAATCATG	2307
QY	2294	acagaaa 2300	
Db	2308	AAAAAA 2314	

RESULT	3	AF331844	2296 bp	mRNA	PRI	06-MAR-2001
LOCUS	AF331844					
DEFINITION	Homo sapiens SOST (SOST) mRNA, complete cds.					
ACCESSION	AF331844					
VERSION	AF331844.1	GI:13236417				
KEYWORDS						
SOURCE	human.					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
AUTHORS	1 (bases 1 to 2296) Balemans,W., Ebeling,M., Patel,N., van Hul,E., Olson,P., Bloszegi,M., Laczka,C., Wuyts,W., van den Ende,J., Willems,P., Paes-Alves,A.F., Hill,S., Bueno,M., Ramos,F.J., Taccoli,P., Dijkers,F.G., Stratakis,C., Lindpaintner,K., Vickers,B., Foerzler,D. and Van Hul,W.					
TITLE	Increased bone density in sclerosteosis is due to the deficiency of a novel secreted protein (SOST)					
JOURNAL	Hum. Mol. Genet. 10 (5), 537-543 (2001)					
PUBMED	11818178					
REFERENCES	2 (bases 1 to 2296) Balemans,W., Ebeling,M., Patel,N., Vickers,B., Foerzler,D. and Van Hul,W.					
TITLE	Direct Submission					
JOURNAL	Submitted (22-DEC-2000) Medical Genetics, University of Antwerp, Universiteitsplein 1, Antwerp B2010, Belgium					
FEATURES	Location/Qualifiers					
SOURCE	1..2296 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="17" /map="17q12-q21" 1..2296 /gene="SOST" 38..679 /gene="SOST" /note="mutations may cause sclerosteosis" /codon_start=1 /product="SOST" /protein_id="AAK16158.1" /db_xref="GI:13236418" /translation="MOLPLALCELVHFAFVVEGOGMOAFKNDATETIPELGEYR EPPELENKTKMRAENGRRPHRPPTKVSQSRLEHPIRYTDDPSARPVME LYCSQCCPRAILRLPAIYGRGMRPPSPDRCPIDPRYRAORQLLCPGEAPRAKTRR LVASCKRLRLRFHNSLKDPEFGAEARPRGKPRCAKNAQALEMAY"					
gene						
CDS						
BE COUNT	576 a	631 c	607 g	482 t		
GIN						
Query Match	98.8%	Score 2273.2	DB 89	Length 2296		
Best Local Similarity	99.8%	Pred. NO. 0				
Matches 2286	Conservative	0	Mismatches 3	Indels 1	Gaps 1	
11	ctactggaagtggtggtccctctctctgtgtgtgtacatgacatccactggtgtgtg					
1	CTACTGGAAGTGGCGTCCCTCTCTCTGTGCTGTACATGCATGCAGCTCCCACTGGCCCTGTG					
71	tctcgtctgctctgtgtacacacacacacgctctcgtgtatgtgaaaggccgaaggtgtgagc					
61	TCTCGTCTGCTCTGTGTACACACACACACCTTCCGTTATGTGAGAGGCCAAGGGGTGGAGGC					
131	gttcaagaatgatgtacacaggaatcatctcccgagctcggagaagtatcccgagctccacc					
121	GTTCAAGAATGATGTCACAGGAATCATCTCCCGAGCTCGAGAGTACCCCGAGGCTCCACCC					
191	ggagcttgagaaacacaaagacatctgaacccggcgagagacgaagagtgagctccacaca					
181	GGAGCTTGAGAACACAAAGACATCTGAACCGGCGGAGAGACGAGAGGGCGGCTCCCAACA					
231	ccctcttgagacaaagacgtgttcgagtatgaagctgtccggagctgtcacttaccgcgtta					
241	CCCCTTTGAGACAAAGACGTGTTCGAGTAGAGCTGCCGCAAGCTGTGACATTTCAACCGGCTA					

OY	311	cygacacgaatggccggatggccgagccgaagccgggtacacgaatctgtgtgtcccgagca	370
Db	301	CGTGAACCCATTGGCCGCTGTCCGACACGCCCAACCCGGTACACAGTGTGTGCTCCGGGCA	360
OY	371	gtgcgagcccgagcgcgctgtctgtcccaagccatcggcccgcgcaagtgtgtgcacataag	430
Db	361	GTCCGGCCCGGCGCGCCCTGTGCTGTCCCAACCCATCTCGGCCGCGGCAAGTGTGTGGACCTAG	420
OY	431	tggagcccgaccttcgctgtacatcccccacccgacacgctgcgagcggtgagaatctgtg	490
Db	421	TGGGCCCGACTTCGCGCTGTCAATCCCGACCGCTTACCGCGCGCAGCGGCTGTGAGTGTGTG	480
OY	491	tcccggtgtgtgagcgcgccgcgcgcgcaaggtgcgctgtgtgtgctcgtgtcaatgtcaa	550
Db	481	TCCCGGTGTGTGAGCGCGCGCGCGCGCAAGGTGTGCGCTGTGTGCTGTGTCAAGTGTCAA	540
OY	551	ggccctcacccgctctccacaacccagtcgagatccaaagatctgggagccgagccgctcg	610
Db	541	GGCCTCACCCGCTCTCCACAAACAGTGTGAGGTCAAGGACTTTGGGAGCGAGCGCGCTCG	600
OY	611	gcccgcagaaaggccgaaagccgcgcgcccgcgcccgagcgccaaagccaaacgagccga	670
Db	601	GCCGCGAAGAGGCGCCGAGAGCCGCGCGCCGCGCCGCGGAGCCCAAAACCAACGAGCCGA	660
OY	671	gctggagaaacgctctactagaaaccccgcccgccctccccaacgagcgcgcccgccgc	730
Db	661	GCTGGAGAAACGCTCTACTAGAACCCCGCCGCGCCCTCCCAACGCGGCGGCGCCCGGGCC	720
OY	731	tgaacccgcgcgcccaattctgtctctgtcggtgtgtgttatgtttalatattatgtt	790
Db	721	TGAACCCGCGGCCCAATTCTGTCTCTGTGCGGTGTGTGTATGTTAATTTCAATTTGT	780
OY	791	aaatgcctcgaaccccaagcgagaggggtctgagaccttcaggccctctgagatcccgagcg	850
Db	781	AAATGCTCTGAACCCCAAGGCGAGGGGGCTAGAACTTCCAGGCGCTTGAGAAATCCGGGGG	840
OY	851	ccgagcaagccccctcaagcccgccagctgtgaggggtcccaacgggacgggaggggaattg	910
Db	841	CCGAGCAAGCCCCCTCTAGCCCGCCCACTGTAGGGGTCCCAAGGGGCGAGGGAGAGGAATTG	900
OY	911	agagtcacagacactgaaacccagcccgccctctggggccgagctacaccttctgtatcc	970
Db	901	AGAGTCAAGACACTGAAACCCAGCCCGCCGCTGTAGGGGTCCCAAGGGGCGAGGGAGAGGAATTG	960
OY	971	caacttcagaggggcagaagaatggaagatlttlcacogccctgggggttttaagggagcggt	1030
Db	961	CACCTTAGAGGAGGCGAAGAAATGGAAGCATTTTTCACCGCCCTGGGGGTTTAAAGGAGACGGT	1020
OY	1031	gtggaggtgaggaagtcaccaagagctgtgttaagaagtttgaataagatltccccccttgac	1090
Db	1021	GTGGAGGTGGGAAGTCCACAGGAGACTGGTTAAGAAATGTGGAATGAATTTCCCTTCACAC	1080
OY	1091	tcgctgcgccatcagaagaacctggaagcggtgcccgagcacaaagctgtgggggcaactgtga	1150
Db	1081	TGCGTGCCTCATGAAAAACCTGTGAGGCGCTGTGCCAGACCAAGAACTGGGGCAACTGTATGA	1140
OY	1151	tgtgtcttcctagtcctgtgctctgcacctaactctgtctgttgaactcttgaactacaacatlc	1210
Db	1141	TGTGTGTTCTAGTCTTGCGCTGTGCACTACTGCTGTGTGAACCTGTGACTACACAAATTC	1200
OY	1211	tccttcggagacctcaattccactttgtlaaatlgaggtgtgaggttgggaatagatctcgg	1270
Db	1201	TCTCTCGGGAACCTCAATTTTCCACTTTTAAAAAGAGGAGGTGAGGTGAATGAAGACTCG	1260
OY	1271	aggagacatttggcatatgatctccaagaagatccagtgaccttlttaatcgggcagaggttga	1330
Db	1261	AGGAGACTTATGTGGCATATGATGATTTCCAAAGAGACTCCAGTGCCTTTGAATGGGCAAGGTGAG	1320
OY	1331	agagagagagagaaagagagagaatgaatgcagttgtcatgtatctcagttgccaaagttcact	1390
Db	1321	AGAAGAGGAGAGAGAGAGGAATGAAATGAAATCAATTTGCAATTTCAATGTGCCAAGGTCACT	1380

Qy	1391	tcggaatcgaaggtgtgtgtgtcctcttcgacagccaaagatataaanaacaaacagaa	1430
Db	1381	TCcGAAATTcAGAGTTGTGTGTGTCTCTTGTGACMcCCAAAGATAAAAACAACAGAA	1440
Qy	1451	aaaaaaagtaaaaggtctatctta tggcgtfacatatttaacgcgtgaacaactccgtgaaga	1510
Db	1441	AAAAAAAGTAAAGAGTCTATTATATGTGGCTGACATATATTACGGCTGACAAACTCTGGAGA	1500
Qy	1511	agcctagtctctcccaagccttggtctcccgagatcgtttgacctccaccctccatct	1570
Db	1501	AGCTATCTGCTTCCACAGCTGGCTTCCCGAGTGTtTGACTACCTCCACCCCTCCACT	1560
Qy	1571	caaaagaaataacatcaltccatttgggttaaaaaagagagtgctcaggtgtgtgtgaag	1630
Db	1561	CAAAAGAAATACATCATCTCCATTGGGGTAAAAAAAGAGAGGGTCCAGAGGTGTGGAGGG	1620
Qy	1631	atagaataacatccgcgcccaactctccaaagaagcagcatccctcccccagccatagcc	1690
Db	1621	ATAGAAATTCATCTCCGCCCAATTCGCCAAGAGCAGCATCTCTCCCGACCAATAGCC	1680
Qy	1691	atgttcttaaaagtcactctccgaagaagatgaagaagttcaagagacatgyccttcagcc	1750
Db	1681	ATGTTTAAAGTCACTCTCGGAAGAGAAAGTGAAGGTTCAAGACACAGCGCTTGACAGC	1740
Qy	1751	ccgagggaagacgcatcaacaaactcagacagcacatccctcttgagaccgcgctc	1810
Db	1741	CCGAGGAGAGCGCATCACAAACTCAGACAGCAGCANCCTTTTGAGACACCGGCGTTC	1800
Qy	1811	tggccaccactcagacacacattctcgtctaaaaaacagctcttactgccttaactgt	1870
Db	1801	TGCCACCACTCCAGGACACATTTCTGCTGTAAANAACGCTTCTTACTGCTTACACATG	1860
Qy	1871	gatggcatccttacciaataaagaatatatttgggggaaaaactcaagtgctgtacata	1930
Db	1861	GATGGCATATCTTACACTAAAGAAATATTATTTGGGGGAAAAACTCAAGTCTGTACATA	1920
Qy	1931	tgctggaanaactcagaagcatatagctgcacacccaanaactcttgaanaactcttc	1990
Db	1921	TGCTGANAAACTGCAGACGATATA CTGCCACCCCAAAATCTTTTGAANAATATTTCC	1979
Qy	1991	agacaacctctcactcttcgtgtagtttllaattgtttaaaaaaaagtttaacaga	2050
Db	1980	AGACAACCTCTTCTCTTCTGTGTAGTTTTTAATTTGTAAAAAAGTTTAAACAGA	2039
Qy	2051	agcacatgacatctgaagccctgcagagctgtctcttcttggaactctccagctg	2110
Db	2040	AGCACATGACATATGAAGAGCTGCAGAGACTGTGCTTTTTTGGCAATTTCTCACGTGG	2099
Qy	2111	gactctgcccaagaatlgaaagttagtgttctttaaaggttaagtlacatatcttacttc	2170
Qy	2100	GACTCTGCCCAAGATGAAGACTGAGTGTTTTAAAGGTTTAAATTTATTTTTC	2159
Qy	2171	tcacttaagtattcttgcnaaagtttctctgttagaagatgacatgttataatitgct	2230
Db	2160	TCACTTAAGTATTATTTATGCAAAAGTTTTCTGTGTGAAGAAAGACAATGTATATTTGCT	2219
Qy	2231	tatgaatlaacagtgtgtctctccagaagtcagagaacatttataaagcaatgact	2290
Db	2220	TATGAATTAACAGTCTGTTCTTCCAGAGTCCAGAGACATTTGTATATTAAGCAATGAATC	2279
Qy	2291	atgacccgaa 2300	
Db	2280	ATGAAAAAA 2289	

[illegible]

ORGANISM	Homo sapiens
ENKARYOTA:	Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
MAMMALIA:	Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 21501)
AUTHORS	Brunkow,M.E., Gardner,J.C., Van Ness,J., Paepel,B.W., Kovacevich,B.R., Prohl,S., Skonier,J.E., Zhao,L., Sabo,P.J., Fu,Y.H., Altsch,R.S., Gilllett,L., Colbert,T., Tacconi,P., Galas,D., Hemsma,H., Beighton,P. and Mulligan,J.T.
TITLE	Bone Dysplasia Sclerosteosis Results from Loss of the SOST Gene Product, a Novel Cysteine Knot-Containing Protein
JOURNAL	Am. J. Hum. Genet. 68 (3), 577-589 (2001)
PubMed	11179006
FEATURES	2 (bases 1 to 21501)
AUTHORS	Brunkow,M.E., Gardner,J.C., Van Ness,J., Paepel,B.W., Kovacevich,B.R., Prohl,S., Skonier,J.E., Zhao,L., Sabo,P.J., Fu,Y.H., Altsch,R.S., Gilllett,L., Colbert,T., Tacconi,P., Galas,D., Hemsma,H., Beighton,P. and Mulligan,J.T.
JOURNAL	Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 1631 220th St. SE, Bothell, WA 98021, USA
FEATURES	Location/Qualifiers
Source	1. 21501
mrna	/organism="Homo sapiens"
cds	/db_xref="taxon:9606"
	/chromosome="17"
	/map="17q21"
	join(<11961..12180,14939..>15360)
	join(11961..12180,14939..15360)
	/codon_start=-1
	/product="sclerostin"
	/protein_id="AK13451.1"
	/db_xref="GI:13161011"
	/translation="MQPLALCLVLVHTAFRVVEGQMAFNKDATEIIPLEGEPYRPPPELENNKTMNRKLENGRPPHPPEPTKSPVSCRELRYVTDFPCSAKPYTEIIVSCGCGARLLPMAIGRGMWRPSGPDRCIPDTRRAVOLLCPGEMAPRRKVMVLVASCCKRLTFPHNUSLEIKDFGEARLPQKGRPPRRASAKNAQAELENNY"
BASE COUNT	5546 a 5259 c 5419 g 5277 t
ORIGIN	
Query Match	89.0%; Score 2047.8; DB 89; Length 21501;
Best Local Similarity	98.9%; Pred. NO. 0;
Matches 2061; Conservative	0; Mismatches 22; Indels 0; Gaps 0;
Oy 219	cgagcgagagaagcgagcgagcgctcccaaccccttgagaccagaagcgtgtccgag 278
Db 14890	CAGGGGTGGCGCAGAGCGGGGGCGACCCCTCAGCGCGCCTCTCTCACAAGACGtGTCGAG 14949
Oy 279	tacagctgtccgcgagagctgtcaacttaccgcgctacgtgtgacgagtgtgcccgttcgcgagcgcc 338
Db 14950	TACAGCTGTCCGCGAGCTGTGACATTACCCGCTACGTAAGACGATGTGGCCGTCCCAACGCGCC 15009
Oy 339	aagcgctaacacgagctgtgtgtcccgccgagctgtgcgcccgcgcgtcgtcccaac 398
Db 15010	AAGCGGCTCACCGAGCTGTGTCTCCGGCCAGTCCGGCCCGGCGCGCTGTGTGCCAAC 15069
Oy 399	gcacatgcgcgcgaggaagatgtgtgtgagactgtgtgagcccaacttcgcgttcatcccgagc 458
Db 15070	GCATTCGGCGCGCGCAGATGTGTGTGGCACCTAATGGGCCCACTTCGCTCAATTCGCCGAC 15129
Oy 459	cgctaccgcgcgacgctgtcagctgtcgtgtgtcccgatgtgtgtagtgcgcgcgcgcgcgcgcgc 518
Db 15130	CGCTACCGCGCGCCGACGCGCTGCAGCTGTGTGTCTCCGGTGTGTAGGCGCGCGCGCGCGCC 15189
Oy 519	aaggtgtccctgtgtgtgcctcgtgtgaagatgtgaagcgccctaccaccccttccacaacacagttg 578
Db 15190	AAGGTGGCGCTGTGTGGCTCTGTGCAAGTGTGCAAGCGCCCTCACCCCTTCCACAACAGTGG 15249
Oy 579	gaagctaaagacttcggagacgagcgcgctcgtgcgcgaagaagcgagacccgcgcgc 638
Db 15250	GAGCTCAAGAGACTTCGGAGACCGAGCGCCCTCTGGCGCGCAAGAGGCGCGAAGCGCGCGCC 15309

QY	639	cgcccccgagcgcccaaaacccaaacgagccgagctctgagaagcctactagaagcccccgc	698
Db	15310	CGCGCCCGAGCGCCCAAGGCCAACAGGCGCAGCTGGAGAACGCTACTTGAAGCCGGCC	15369
QY	699	ggcgcctccccaacgagcgagcgcccgcccgaaaccgagcccaatlctcctc	758
Db	15370	GGCGCCCTCCCAACCGCGCGGGCCCCCGCCTTAACCGCGGCCCAATTCGTGCTC	15429
QY	759	tgcgcgtgcttatgttatattcaatctgtaatgcctgcgaaccgaagcgagtgagct	818
Db	15430	TGCGCGGTGTATGTTATTTATTTCAATTGTAATGCTGTCAACCCAGGCGAGGGGCT	15489
QY	819	gagaccttccagggccctgagaaatcccgagcgcccggaagggcccccctagcccgcaac	878
Db	15490	GAGACCTTCAAGGCGCTGAGGAAATCCCGGCGCGGCAAGGCCCTCTAGCCCGCAAC	15549
QY	879	tgaagggctcccaagggcgagggaggaattgaagatcaacagacatgaagccagagcc	938
Db	15550	TGAGGGGTCCCAAGGGGCGAGGAGGAATTAAGAGTACAGACACTGAGCCACGAGCC	15609
QY	939	ccgcctcttggggcgagcctaacttctgctccaaactcaagggggagcaaatgagaaga	998
Db	15610	CCGCGTCTGGGGCGGCTACCTTCTGCTGCTCCACTTCAGAGGAGGCAAAATGGAAGA	15669
QY	999	ttttcaacggccctgggggttttaaggaagcggtgtggaagtgggaagatccaagatggt	1058
Db	15670	TTTTTACCGCCCTGGGGGTTTTTAAGGAGGGGTGGGAGTGGGAATCCAGGAGCTGGT	15729
QY	1059	taagaaagtgtgaataatctccctctgcacctgcgtcccatcaagaagccctgaagcgt	1118
Db	15730	TAAAGAAAGTTGGATAGATATTCCTCCCTGCACCTGCTGCCTACAGAAACCTGAGCGT	15789
QY	1119	gcccagaagcaagaagcttgggggcaactctagatgtggtttctagctctgctgcact	1178
Db	15790	GCCAGAGCAACAAGACTGGGGGCMACTGATGATGTTCTATCTCTGGCTCTGGCACT	15849
QY	1179	aactctgtgtgaacttgaactaacaattctcttcggagactcaattccacttgt	1238
Db	15850	AACCTGCTGTGAACCTTAACTTAACACAAATTCCTCTGGGACCTCAATTTCCACTTGT	15909
QY	1239	aaaaatggagtggaagtggaatagatctcgaagagacatctgcatatgatccaag	1298
Db	15910	AAAATGAGGGGTGAGGTGGGAATAGAGATCTCGAGGAGACTTGTGGCATATGATTTCCAAAG	15969
QY	1299	actccaatgctttbaatgggacagagtggagaagagaagagaagaagaagaatga	1358
Db	15970	ACTCAATGCCCTTTGAATGGCGACAGGTGAGAGAGAGAGAAAGAGAGAGATGA	16029
QY	1359	tgcagctgcatctgaatcagltgcgaagtlcaacttccagaatcagaatgtgtagtctct	1418
Db	16030	TGCGATTTGATTTATTTACGTGCGCAAGGTCACTCCAGAAATTCAGAGTTGTGATGGCTCT	16089
QY	1419	tctacacgcaaaagatgaaacaaacaaagaaaaaaagaaagagatctattatgct	1478
Db	16090	TCTACACGCCAAAGATGAAAAACAAACAGAAAAAAGTAAGAGTCTATTATATGGCT	16149
QY	1479	gaacatttcacgctgaacaaactctgaaaaaagatagtgcttcccacgcagcttcc	1538
Db	16150	GACCTATTTACGGCTGACAACTCCTGGAAGAGTATGCTGCTTCCAGCTGCGCTTCC	16209
QY	1539	ccgagatgttgcttacctccaccctccatctcaagaatatcaatcalccatltgggta	1598
Db	16210	CCGATTTTGGCTACTCTCACCCCTCCATCTCAAAAGAAATATCATTCATTTGGGGTA	16266
QY	1599	gaaagaagagggctccagagtggtggaagatagaatcacatccgcccacaattcc	1658
Db	16270	GAAAGAGAGGGGTCGAGGGGTGTGGAGGATGGAATACATCCGCCCAACTTCCC	16329
QY	1659	aaagaagagatccctcccccgaaccatagacatgttttaagctcaacttccgaagaaga	1718
Db	16330	AAAAGACAGATCCCTCCCGACACCATATGCTATTTTAAAGTCACTTTCGAGAGGA	16388
QY	1719	gtgaagagttcaagaaactgctcttgagcgcccgagggagcagcatcacaactcaca	1778

	Db	16390	GTCAGACGTTCAAGCAGCACTGGCGTTCGAGGCCGAGAGGAGACCATCATCAAACTCACCA	16449
Oy	1779	gaccagcacalcccttlttgagacacggccttcgtcccaactcaaggaaactttctgc		1838
Db	16450	GACCAGCACATCCCTTTTAGACACCGGCTTCGTGCACAACATCAGGACACATTTCTGC		16509
Oy	1839	ctagaacaacgcctctctacgtccttaacatlygaatlygcatactcttaactaaaagaat		1898
Db	16510	CTAGAACAACGCTTWTACTGCTTACATGATGATGATGCATATCTTAACATAAAGAATAT		16569
Oy	1899	tattgggggaaaaaacacaaagfycctgcatatgctgagaacatcgacagcataatagt		1958
Db	16570	TATTGGGGGAAAAAACATCAAGTGCTGTACATNTGTGAGAAAACGACAGACATATAGCT		16629
Oy	1959	gccacccaanaactcctlttgaaatcatctccagacaacactctaacttctgtagtgtt		2018
Db	16630	GCCACCCCAAATACTTTTTGAATAATCATTTCCAGACAACCTCTTACTTCTGTGTAGTTT		16689
Oy	2019	ttaatctttaaaaaaaaaaagtltaaacaagaagacatgacatgatlaaagacctgcaga		2078
Db	16690	TTAATTTGTAAAAAATAAAGTTTAAACAGAGACATGCATATATGAAGCCTGCAGGA		16749
Oy	2079	ctgagtcgttttttggcaattcttcocagtgtagcttgtccacaagaatgaagtagtg		2138
Db	16750	CTGAGTCGTTTTTTTGGCAATCTTCCACGTGGACCTTGTCCACAGATATGAAGTAGTGG		16809
Oy	2139	ttttaaagsgtlaagtlacatatatttatlctcoactlaagtlattatgacaaagtltt		2198
Db	16810	TTTTTAAAGGATTAAGTTACATATATTTATTTCTCCATTAACTTATTTATCCAAAAGTTT		16869
Oy	2199	tctgttgagaaatgacaaatgtaaatatgtccttiaatgaaataacgctgtcttccagag		2258
Db	16870	TCTTGTGAGAAATGACAAATGTTAATTTCTTTATGAATTAACAGTCTGTTTCCAGAG		16929
Oy	2259	tcscagacatgttataataagaacaataatcatatgacccgaag		2301
Db	16930	TCCAGACACATTTGTTATATTAAGCATGATGATATGACCGAAG		16972
RESULT	5			
LOCUS	AC003098/c	94752 bp	DNA	PRI
DEFINITION	Homo sapiens chromosome 17, clone HRPC905M1, complete sequence.			29-JAN-1998
ACCESSION	AC003098			
VERSION	AC003098.1	GI:2822155		
KEYWORDS	HTG;			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 94752)			
AUTHORS	Birren,B., Fasmann,K., McKernan,K., Nusbaum,C., Richardson,P. and Lander,E.			
TITLE	Homo sapiens chromosome 17, clone HRPC905M1			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 94752)			
AUTHORS	Birren,B., Fasmann,K., McKernan,K., Munro,C., Nusbaum,C., Richardson,P., Lander,E., Baldwin,J., Barra,N., Cantu,C., Coake,P., Daly,M.J., Devon,K., Dewar,K., Dukette,B., Forrest,C., Gage,D., Genschelmer,S., Gerajery,K., Gilmarlin,T., Hagos,B., Halphen,I., Harris,K., Howland,J.C., Huang,J., Hul-B., Jacotot,L., Kirby,A., Lane,M., Mackenzie,J., Marguis,N., Mcdermott,J., Molla,M., Morrow,J., Nachman,A., Naylor,J., O'Connor,T., Olotu,A., Peterson,K., Roberts,D., Rollins,G., Sarnaik,A., Shiu,P., Shyam,R., Stillell,J., Stone,C., Strickland,C., Sydney,K., Tang,L., Zemseva,I. and Zody,M.			
TITLE	Direct Submission			
JOURNAL	Submitted(10-NOV-1997) Whitehead Institute/MIT Center for Genome Research, <del>3</del> Chenieries Street, Cambridge, MA 02141, USA			
REFERENCE	3 (bases 1 to 94752)			
AUTHORS	Birren,B., Fasmann,K., McKernan,K., Nusbaum,C., Richardson,P., Lander,E., Allen,N., Baker,J., Baldwin,J., Barra,N., Beckerly,R.,			

Boutwell,C., Byrne,S., Cantu,C., Castle,A., Cooke,P., Daly,M.J.,  
 Depayre,E., Devon,K., Dewar,K., Durette,B., Etemadi,S.,  
 Ferreira,P., Forrest,C., Gage,D., Gardyna,S., Gensheimer,S.,  
 Geraghty,K., Gilmartin,T., Gray,D., Haggos,B., Harris,K.,  
 Horton,L., Howland,J.C., Hul.L., Jacotot,L., Litton,L.,  
 Mackenzie,J., Marquis,N., McEwan,P., McGuck,A., Meldrum,J.,  
 Molla,M., Morris,W., Morrow,J., Nachman,A., Naylor,J., O'Connor,T.,  
 Pavlin,B., Peterson,K., Ranganath,S., Riley,R., Roberts,D.,  
 Rollins,G., Rossello,R., Roy,A., Shyam,R., Soohoo,S., Stillwell,J.,  
 Stone,C., Strickland,C., Sydney,K., Tang,L., Vassiliev,H., Vo,A.,  
 Wagner,A., Wheeler,J., Wu,Y., Ye,W.J., Zemtseva,I., Zhao,J. and  
 Zody,M.

TITLE Direct Submission  
 JOURNAL Submitted (29-JAN-1998) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 COMMENT On Jan 29, 1998 this sequence version replaced gi:2665504.  
 The Staden databases, finishing information, and all  
 chromatographic files used in the assembly of this clone are  
 available from our anonymous ftp site.

All repeats were identified using RepeatMasker: Smit, A.F.A. &  
 Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>.

FEATURES  
 source  
 1..94752  
 Location/Qualifiers

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="HRC905N1"  
 /clone\_11b="Peter de Jong/ human PAC library"  
 /map="17"  
 /chromosome="17"  
 86..439  
 /rpt\_family="LINE2"  
 complement(972..1262)  
 /rpt\_family="AlusX"  
 complement(1884..1912)  
 /rpt\_family="(CA)n"  
 1912..1939  
 /rpt\_family="(GA)n"  
 2633..2681  
 /rpt\_family="MIR"  
 2701..3002  
 /rpt\_family="AlusX"  
 complement(3874..4008)  
 /rpt\_family="LINE2"  
 4513..4543  
 /rpt\_family="AT-rich"  
 complement(5202..5337)  
 /rpt\_family="(GA)n"  
 complement(5310..5406)  
 /rpt\_family="MIR"  
 complement(5829..5868)  
 /rpt\_family="GC-rich"  
 6485..6592  
 /rpt\_family="MIR"  
 complement(6973..7029)  
 /rpt\_family="MIR"  
 complement(10402..10426)  
 /rpt\_family="GC-rich"  
 complement(11421..11496)  
 /rpt\_family="LINE"  
 complement(11505..11804)  
 /rpt\_family="AlusX"  
 11811..11879  
 /rpt\_family="AT-rich"  
 11880..12093  
 /rpt\_family="AluYb"  
 complement(12104..12476)  
 /rpt\_family="LINE2"  
 12478..12773  
 /rpt\_family="AlusX"  
 complement(12775..12917)  
 /rpt\_family="LINE2"

repeat\_region complement(13088..13510)  
 /rpt\_family="MLTIC"  
 repeat\_region complement(13544..13847)  
 /rpt\_family="AlusX"  
 repeat\_region complement(14110..14776)  
 /rpt\_family="LINE1"  
 repeat\_region complement(14777..15077)  
 /rpt\_family="AlusX"  
 repeat\_region complement(15079..15282)  
 /rpt\_family="LINE1"  
 repeat\_region 15954..16024  
 /rpt\_family="AT-rich"  
 repeat\_region 18031..18178  
 /rpt\_family="MIR"  
 repeat\_region 20398..20769  
 /rpt\_family="LINE"  
 repeat\_region 20893..21190  
 /rpt\_family="Alusg"  
 complement(21275..21473)  
 /rpt\_family="Alusg"  
 complement(21485..21556)  
 /rpt\_family="(CA)n"  
 repeat\_region 21567..21906  
 /rpt\_family="MER2"  
 repeat\_region 21907..21952  
 /rpt\_family="MLT2CB"  
 repeat\_region 21974..22115  
 /rpt\_family="MIR"  
 repeat\_region 22238..22540  
 /rpt\_family="AlusX"  
 repeat\_region 22577..22633  
 /rpt\_family="LINE2"  
 repeat\_region complement(23190..23270)  
 /rpt\_family="LINE2"  
 24194..24324  
 /rpt\_family="FLAM\_A"  
 24328..24643  
 /rpt\_family="Alusg"  
 25542..25698  
 /rpt\_family="LINE2"  
 25805..26001  
 /rpt\_family="MIR"  
 repeat\_region 26060..26174  
 /rpt\_family="LINE2"  
 repeat\_region 26343..26649  
 /rpt\_family="AlusX"  
 complement(27491..27783)  
 /rpt\_family="Alusg"  
 27862..27906  
 /rpt\_family="LINE2"  
 repeat\_region 28068..28114  
 /rpt\_family="MIR"  
 repeat\_region complement(28329..28360)  
 /rpt\_family="AT-rich"  
 28371..28652  
 /rpt\_family="Alusg"  
 complement(29498..29604)  
 /rpt\_family="GC-rich"  
 30074..30181  
 /rpt\_family="MIR"  
 complement(30181..30249)  
 /rpt\_family="LINE2"  
 repeat\_region 30444..30517  
 /rpt\_family="MIR"  
 complement(31852..32135)  
 /rpt\_family="Alusg"  
 complement(32675..33003)  
 /rpt\_family="MLT1A1"  
 complement(33042..33082)  
 /rpt\_family="LINE2"  
 33363..33459  
 /rpt\_family="MER5B"  
 repeat\_region complement(35475..35524)

[illegible]

Db 4600 GCCACCCAAAATCTTTTGAATCATTTCCAGAACCTCTTACTTCTGTAGTTT 4541  
QY 2019 ttatgtttaaaaaaaagttttaaagaagacacatgatattgaagcctgcagga 2078  
Db 4540 TTATATGTTTAAAAAAGTTTAAACAGAGACATGACATATGAAAGCCTGCACGA 4481  
QY 2079 ctgttcgttttttggcaattcttccacgtgagctgtgccagaagaagtaagtcg 2138  
Db 4480 CTGTGCTTTTGTGGCAATTTCTCCAGCTGGAGCTTGTCCACAGAAATGAAGTAGTG 4421  
QY 2139 tttttaagaagtaagttacatattatttttccactaagttattttgcgaagttt 2198  
Db 4420 TTTTAAAGAGTTAACTTACATATTTATTTCTCACTTAAGTTATTAAGCAAAAGTTT 4361  
QY 2199 tctgttagaagatgaacatgttaataatgtcttaacgaacagtcgtcttccagag 2258  
Db 4360 TCTTGTAGAGAAATGACATGTTAAATTTGCTTTATGAATTAACACTGTTCTTCCAGAG 4301  
QY 2259 ttcagagacatgttataaagaacatgaatcatgaccgaaag 2301  
Db 4300 TCCAGAGACATTTTATTAAGACAAATGATCATGACCGAAAG 4258

AC055813 151780 bp DNA HTG 22-JUL-2000  
LOCUS Homo sapiens chromosome 17 clone RP11-209M4 map 17, WORKING DRAFT  
DEFINITION AC055813  
AC055813  
AC055813 GI:9369521  
VERSION HTG: HTGS\_PHASE1, HTGS\_DRAFT.  
KEYWORDS human.  
SOURCE Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
ORGANISM Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 151780)  
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
TITLE Homo sapiens chromosome 17, clone RP11-209M4  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 151780)  
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
Anderson,S., Baldwin,J., Barne,N., Bastien,V., Beda,F.,  
Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G.,  
Campiano,A., Castle,A., Choquel,Y., Colangelo,M., Collins,S.,  
Collamore,A., Cooke,P., Dextrallano,K., Dewar,K., Diaz,J.S.,  
Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,  
Galagan,J., Gardyna,S., Glade,S., Goyette,M., Graham,L.,  
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,  
Klein,J., Lacroque,K., Lamazares,R., Jones,C., Kann,L., Karatas,A.,  
Klewin,J., Larocque,K., Lamazares,R., Landers,T., Lenoczky,J.,  
Levine,R., Lien,C., Liu,G., Locke,K., Macdonald,P., Margolis,N.,  
McCarthy,M., McEwan,P., McGuck,A., McKernan,K., McNeelers,R.,  
Meltrin,J., Meneses,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,  
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,  
O'Neil,D., Olyar,T.M., Oliver,J., Peterson,K., Pierre,N.,  
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,  
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,  
Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
Testaye-S., Theodore,J., Tirrell,A., Travers,M., Triggiani,J.,  
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,  
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

TITLE Direct Submission  
JOURNAL Submitted (18-APR-2000) Whitehead Institute/MIT Center for Genome  
COMMENT Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jul 22, 2000 this sequence version replaced gi:8920653.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIRB  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu

----- Project Information  
Center project name: 17949  
Center clone name: 209.M.4  
----- Summary Statistics  
Sequencing vector: M13; M77815; 94% of reads  
Sequencing vector: Plasmid; n/a; 0.1% of reads  
6.1588306320908Chemistry: Dye-terminator Big Dye; 100% of  
reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 139835 bases at least Q40  
Consensus quality: 145751 bases at least Q30  
Consensus quality: 148014 bases at least Q20  
Insert size: 155000; agarose-1p  
Insert size: 149480; sum-of-ctgifs  
Quality coverage: 3.9 in Q20 bases; agarose-1p  
Quality cov.  
NOTE: This is a 'working draft' sequence. It currently  
\* consists of 24 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1 1283: contig of 1283 bp in length  
\* 1284 1383: gap of 100 bp  
\* 1384 2931: contig of 1548 bp in length  
\* 2932 3031: gap of 100 bp  
\* 3032 4160: contig of 1129 bp in length  
\* 4161 4260: gap of 100 bp  
\* 4261 6377: contig of 2117 bp in length  
\* 6378 6477: gap of 100 bp  
\* 6478 8584: contig of 2107 bp in length  
\* 8585 8684: gap of 100 bp  
\* 8685 10824: contig of 2140 bp in length  
\* 10825 10924: gap of 100 bp  
\* 10925 13710: contig of 2786 bp in length  
\* 13711 13810: gap of 100 bp  
\* 13811 17466: contig of 3656 bp in length  
\* 17467 17566: gap of 100 bp  
\* 17567 19747: contig of 2181 bp in length  
\* 19748 19847: gap of 100 bp  
\* 19848 22297: contig of 3150 bp in length  
\* 22998 23097: gap of 100 bp  
\* 23098 26840: contig of 3743 bp in length  
\* 26841 26940: gap of 100 bp  
\* 26941 32331: contig of 5391 bp in length  
\* 32332 32431: gap of 100 bp  
\* 32432 37589: contig of 5158 bp in length  
\* 37590 37689: gap of 100 bp  
\* 37690 43814: contig of 6125 bp in length  
\* 43815 43914: gap of 100 bp  
\* 43915 49593: contig of 5681 bp in length  
\* 49596 49695: gap of 100 bp  
\* 49696 56982: contig of 7287 bp in length  
\* 56983 57082: gap of 100 bp  
\* 57083 63903: contig of 6821 bp in length  
\* 63904 64003: gap of 100 bp  
\* 64004 73106: contig of 9103 bp in length  
\* 73107 73206: gap of 100 bp  
\* 73207 81472: contig of 8266 bp in length  
\* 81473 81572: gap of 100 bp  
\* 81573 94653: contig of 13081 bp in length  
\* 94654 94753: gap of 100 bp  
\* 94754 106759: contig of 12006 bp in length  
\* 106760 106859: gap of 100 bp  
\* 106860 118633: contig of 11774 bp in length  
\* 118634 118733: gap of 100 bp  
\* 118734 135004: contig of 16271 bp in length  
\* 135005 135104: gap of 100 bp  
\* 135105 151780: contig of 16676 bp in length.  
Location/Qualifiers  
1..151780

FEATURES  
SOURCE





Db 113251 TCTGACAGCCAAAGATGCAAAAAACAACAGAAAAAAGTAAAGACTTATTATGGCT 113310  
QY 1479 gacataattacggcgcagacaactccggaagaagctatgctcctccagcgcgtctcc 1538  
Db 113311 GACATATTTACGGCTCACAACACTCTGGAGAGAGTATGCTTCCAGCTGGCTCC 113370  
QY 1539 ccgagctgttggtacccaccctccatctcaagaagaatacatcatccatgaggtta 1598  
Db 113371 CCGGATCTTTGGCTACCTCACCCTCCATCTCAAGAAATTAATCATCTCATTTGGGGA 113430  
QY 1599 gaaaagagaaggtcccgaggtgtgfggaaggaatagaatcaatccgcccaactccc 1658  
Db 113431 GAAAGAGAGAGGCTCCGAGGGGTGGAGGATGAAATCAATCCGCCCAACTTCCC 113490  
QY 1659 aaagagcaagatccctcccccagccatagccatgttttaagtcacactccgaagaa 1718  
Db 113491 AAAGAGAGATCCCTCCCGACCCATAGCCATGTTTAAAGTCACCTTCCGAGAGAA 113550  
QY 1719 gtgaagagttcagaagacatgctctgcaagccgaagagcagccatcaacaactaca 1778  
Db 113551 GTGAAGGTTCAAGACACTGGCTTGCAGGCCCGAGGACACCATCAACAACTACA 113610  
1779 gacacagacatccctcttgagacacgcctctgcccacacacacagcagacatcttctg 1838  
Db 113611 GACACAGACATCCCTTTTGAAGACCGCCTTCTGCCACACATCAGGACACATTTCTGC 113670  
QY 1839 ctgaaagacgcctcttactgctcttacaatgtatgagatcttacaactaaagaatat 1898  
Db 113671 CTGAAAGAGCTTCTTACTGCTTACATGATGATGATGATGATGATGATGATGATGAT 113730  
QY 1899 tacttgggggaaaaaacatcacagctgctgatacatgctgagaacacgcagagcagataatagc 1958  
Db 113731 TATTGGGGGAAAAACTACAAAGTGTACATATGCTGAGAAATGACAGACATATAT-CT 113789  
QY 1959 gccacccaataacacttttgaaatcatcttccagaacaacacttactcttgggtggtt 2018  
Db 113790 GCCACCCAAAATCTTTTGAAGATCATTTCCAGACAACTCTTACTTCTGATGATGAT 113849  
QY 2019 ttaactgttaaaaaaaaagtittaaacgaagacacatgacataagaagcctgcagga 2078  
Db 113850 TTAATGCTTAAAAAAAAGTTTAAACAGAGCATGATGATGATGATGATGATGATGATGAT 113909  
QY 2079 ctgagcttttttggcaatcttccacgtggacttccacaagaataaagatgagtg 2138  
Db 113910 CTGGGTGTTTTTTTGGCAATTTTCCACGTGGGACTTCTCCAGAAATGAAGTACTGG 113969  
QY 2139 tttttaagaagttaaagttacatatatttctcaactaaagttatttaagcaaaagtttt 2198  
Db 113970 TTTTAAAGAGTTAACTATGATTTATTTTCTCACTTAAAGTTATTTCAAAAGTTTT 114029  
2199 tctgttagagaatgacaatgataatctgcttatagaatgaaatgacgtctgcttccagag 2258  
Db 114030 TCTTTAGAGAAATGACAATGTTAATGCTTTATGATTAACAGTCTGTTCTCCAGAG 114089  
QY 2259 tccaagacatgttaataaagaacatgaatcagcaggaag 2301  
Db 114090 TCCAGAGACATTGTTATTAAGACATGATGATGACCGAAG 114132  
RESULT 7  
AC073954 177744 bp DNA HNG 17-NOV-2000  
LOCUS AC073954/c Homo sapiens chromosome 17 clone RP11-147C10, WORKING DRAFT  
DEFINITION SEQUENCE, 34 unordered pieces.  
ACCESSION AC073954  
VERSION AC073954.3 GI:10998993  
KEYWORDS HNG, HNGS\_PHASE1, HNGS\_DRAFT.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 177744)  
AUTHORS Waterston, R.H.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

The sequence of Homo sapiens clone  
Unpublished  
2 (bases 1 to 177744)  
Waterston, R.H.  
Direct Submission  
Submitted (07-JUL-2000) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
On Oct 25, 2000 this sequence version replaced gi:1986849.

## COMMENT

----- Genome Center -----  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: http://genome.wustl.edu/gsc/index.shtml  
----- Project Information -----  
Center project name: H.NH0147C10  
----- Summary Statistics -----  
Sequencing vector: M13, 100%  
Sequencing vector: Plasmid, 0%  
Chemistry: Dye-Primer ET; 100% of reads  
Chemistry: Dye-terminator Big Dye; 0% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 155679 bases at least Q40  
Consensus quality: 163329 bases at least Q30  
Consensus quality: 165780 bases at least Q20  
Insert size: 18500; agarose-fp  
Insert size: 183559; sum-of-contigs  
Quality coverage: 3.29 in Q20 bases; agarose-fp  
Quality coverage: 3.43 in Q20 bases; sum-of-contigs  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 34 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1  
\* 1248 1247: contig of 1247 bp in length  
\* 1348 1347: gap of unknown length  
\* 1348 2762: contig of 1415 bp in length  
\* 2763 2862: gap of unknown length  
\* 2863 4773: contig of 1911 bp in length  
\* 4774 4873: gap of unknown length  
\* 4874 6326: contig of 1453 bp in length  
\* 6327 6426: gap of unknown length  
\* 6427 8673: contig of 2447 bp in length  
\* 8674 8973: gap of unknown length  
\* 8974 10783: contig of 1810 bp in length  
\* 10784 10883: gap of unknown length  
\* 10884 12763: contig of 1880 bp in length  
\* 12764 12863: gap of unknown length  
\* 12864 15381: contig of 2518 bp in length  
\* 15382 15481: gap of unknown length  
\* 15482 17365: contig of 1884 bp in length  
\* 17366 17465: gap of unknown length  
\* 17466 19992: contig of 2527 bp in length  
\* 19993 20092: gap of unknown length  
\* 20093 23512: contig of 3420 bp in length  
\* 23513 23612: gap of unknown length  
\* 23613 25699: contig of 2087 bp in length  
\* 25700 28333: gap of unknown length  
\* 28334 28433: contig of 2534 bp in length  
\* 28434 32027: gap of unknown length  
\* 32028 32127: gap of 3594 bp in length  
\* 32128 36115: contig of 3988 bp in length  
\* 36116 36215: gap of unknown length  
\* 36216 39882: contig of 3667 bp in length  
\* 39883 39982: gap of unknown length  
\* 39983 44009: contig of 4027 bp in length  
\* 44010 44109: gap of unknown length  
\* 44110 47829: contig of 3720 bp in length

*	47830	47929:	gap of unknown length
*	47930	53768:	contig of 5839 bp in length
*	53769	53868:	gap of unknown length
*	53869	60215:	contig of 6347 bp in length
*	60216	60315:	gap of unknown length
*	60316	66477:	contig of 6162 bp in length
*	66478	66577:	gap of unknown length
*	66578	71476:	contig of 4899 bp in length
*	71477	71576:	gap of unknown length
*	71577	78683:	contig of 7307 bp in length
*	78684	78984:	gap of unknown length
*	78984	85952:	contig of 6965 bp in length
*	85953	86052:	gap of unknown length
*	86053	93510:	contig of 7458 bp in length
*	93511	93610:	gap of unknown length
*	93611	100718:	contig of 7108 bp in length
*	100719	100818:	gap of unknown length
*	100819	1009795:	contig of 8977 bp in length
*	100980	109895:	gap of unknown length
*	109896	117930:	contig of 7895 bp in length
*	117991	117890:	gap of unknown length
*	117891	126044:	contig of 8154 bp in length
*	126045	126444:	gap of unknown length
*	126445	134442:	contig of 8298 bp in length
*	134446	134542:	gap of unknown length
*	134543	145663:	contig of 11121 bp in length
*	145664	145763:	gap of unknown length
*	145764	157189:	contig of 11426 bp in length
*	157190	157289:	gap of unknown length
*	157290	160622:	contig of 9533 bp in length
*	160623	166922:	gap of unknown length
*	166923	177944:	contig of 10822 bp in length

UNES	location/Qualifiers
source	1..177744
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="17"
	/clone="RP11-147C10"
misc_feature	1..1247
	/note="assembly_name:Contig26"
misc_feature	1348..2762
	/note="assembly_name:Contig31"
misc_feature	2863..4773
	/note="assembly_name:Contig34"
misc_feature	4874..6326
	/note="assembly_name:Contig35"
misc_feature	6427..8873
	/note="assembly_name:Contig36"
misc_feature	8974..10783
	/note="assembly_name:Contig37"
misc_feature	10884..12763
	/note="assembly_name:Contig38"
misc_feature	12864..15381
	/note="assembly_name:Contig39"
misc_feature	15482..17365
	/note="assembly_name:Contig40"
misc_feature	17466..19992
	/note="assembly_name:Contig41"
misc_feature	20093..23512
	/note="assembly_name:Contig42"
misc_feature	23613..25699
	/note="assembly_name:Contig43"
misc_feature	25800..28333
	/note="assembly_name:Contig44"
misc_feature	28434..32027
	/note="assembly_name:Contig45"
misc_feature	32128..36115
	/note="assembly_name:Contig46"
misc_feature	36216..39682
	/note="assembly_name:Contig47"
misc_feature	39983..44009
	/note="assembly_name:Contig48"
misc_feature	44110..47829

	misc_feature	/note="assembly_name:Contig49"	47930..53768	
		/note="assembly_name:Contig50"	53869..60215	
	misc_feature	/note="assembly_name:Contig51"	60316..66477	
	misc_feature	/note="assembly_name:Contig52"	66578..71476	
	misc_feature	/note="assembly_name:Contig53"	71577..78883	
	misc_feature	/note="assembly_name:Contig54"	78984..85952	
	misc_feature	/note="assembly_name:Contig55"	clone.end:77	
		vector.side:left"	86053..93510	
	misc_feature	/note="assembly_name:Contig56"	93611..100718	
	misc_feature	/note="assembly_name:Contig57"	100819..109795	
	misc_feature	/note="assembly_name:Contig58"	109896..117790	
	misc_feature	/note="assembly_name:Contig59"	117891..126044	
	misc_feature	/note="assembly_name:Contig60"	126145..134442	
	misc_feature	/note="assembly_name:Contig61"	134543..145663	
	misc_feature	/note="assembly_name:Contig62"	145764..157189	
	misc_feature	/note="assembly_name:Contig63"	157290..166822	
	misc_feature	/note="assembly_name:Contig64"	166923..177744	
	misc_feature	/note="assembly_name:Contig65"		
BASE COUNT	44147 a	43012 c	43431 g	43859 t 3315 others
ORIGIN				
Query Match	46.3%;	Score 1066.4;	DB 75;	Length 177744;
Best Local Similarity	97.0%;	Pred. No. 1.7e-195;		
Matches 1129;	Conservative	0;	Mismatches 31;	Indels 4;
				Gaps 4;
QY	1142 aactgtatgctgttct-agtctgctctgcac-taactgcgtctaaacttga	1199		
DB	93369 AGCTGAGATGTTGTGTCTGATGTTCTGTGGCCCACTTACTTGTGGAAATCCGTGA	93310		
QY	1200 ctacacattctccttcggacctcaattcca-ctttgtaaaatgaggtgtgagttgag	1258		
DB	93309 CTACACAAGTGTGTGGGGGACCTCATATGTCATCTGTGTAATAATGATGGTGGAGTGGG	93250		
QY	1259 aatgagatctcgagagagactattggcatalgtatccaagagactccagtgcctttgaag	1318		
DB	93249 AATAGGATCTTCGAGAGACTATTGGCATTTGATGCTCAAGACCTCAGTGCCTTTGATG	93190		
QY	1319 ggcagagtgagagagagagagaaagagagaagataaagtcagttgcattgattcaagt	1378		
DB	93189 GCGAGAGGTGAGAGAGAGAGAGAAAGAGAGAGATGAATGACATTGCAATTCAGT	93130		
QY	1379 gccaaagtcacttcagaatt-caaggtgtgatatctctcttcgcagagccaaagatgag	1437		
DB	93129 GCCAAGGTCACCTTCCAAATTTNAGAGTTGTATGCTCTCTTTCACAGCCAAAGATGAA	93070		
QY	1438 aaacaacagaaaaaaaagtaaaagtcataattatgctctgacataltaacgctgaca	1497		
DB	93069 AAACAAACAGAAAAAAAAGCTAAAGAGTCTATTATGCTGACATATTTCGGCTGACA	93010		
QY	1498 aactcttggaagaagcgtatgctctccagctggtctcccgatgttggtactac	1557		
DB	93009 AACTCTTGAGAGAGCTATGCTCTTCCAGGCTGGCTTCCCGGATGTTTGGCTTACTC	92950		
QY	1558 caaccctcatctcaaaagaaataacataccattggtggtagaagagagaggtccga	1617		

Db	Accession	Organism	Gene	Protein	Length	Score	E-value
Db	92349	CACCCCTCCATCTCAAGAAGAAATACATCATTCATTCATTGGGTAGAAAAGAGAGAGGTCGAG			92890		
Qy	1618	ggttggtggagggatgaatacaatccgccccaaattcccaagaagagatccctccc			1677		
Db	92889	GGTGTGGGAGGATGGAATTCATCTCCGCCCAACTTCGCCAAGAGAGCATCCCTCC			92830		
Qy	1678	ccgagcccatgacatgltttaaagtaacctccgaagaagtgaaggttccaagagcac			1737		
Db	92829	CCGACCATATGSCATCTTTTAAAGTCACCTTCGGAAGAGAGTGAAGAGTTCAGAGAC			92770		
Qy	1738	tggaccttgagcccgcaaggagagcagccatcaaaactcaagaccagaaatcccttgg			1797		
Db	92769	TGGCTTTGCAAGGCCCGAGGAGAGCGCATCAAACTCAAGACACGACATCCCTTTG			92710		
Qy	1798	agaaacgcgccttctgcccacacactcaagagacatcttcgcctagaaaacagctcttac			1857		
Db	92709	AGACACGCGCTTGTGCCCAACACTCAGACACATTTTCGCTGTGAAACAGCTTTTAC			92650		
Qy	1858	tgcctcttaacatgtagatggatctcttaacctaagaagaaatattatgggggaaaaactaca			1917		
Db	92649	TGCCTTACATGTGATGGCATATCTTACATCAAAAGAAATATATATGGGGAAAAACTACA			92590		
Qy	1918	agtcgctgacatatgtctgagaaactgacagagataatagctgcccacccaaaatccttct			1977		
Db	92589	AGTCTGTACATATGCTGAGAAACTCAGACGATATATCTGCCACCTCAAAATCTTTT			92530		
Qy	1978	gaatacatcttcagaacaactcttaactctgctgctgaattttaaattgtaaaaaaaa			2037		
Db	92539	GAAATATATTTCCAGACAACTCTTACTTCTGTGTAGTTTATATGTTAAATAAAAA			92470		
Qy	2038	agttttaaacagaagacacatgacataatgaagaagcctgacagactgcttctttgacaa			2097		
Db	92469	AGTTTAAACAGAGACATATGACATATGAAAACCTCGACGACTGCTGTTTTTGGCAA			92410		
Qy	2098	tctctcaagctggagactcttcacaagaatgaagaatgaatgattttaaagaagtttaacta			2157		
Db	92409	TTCTTCACAGCTGGGACTTGTCCACAAAGATGAAGTAATAGTGGTTTTTAAAGAGTTAACTTA			92350		
Qy	2158	catatttattcttcacttaagatattattatgcaaaagtttctctgtagagaatgacat			2217		
Db	92349	CATATATATATTTCTCACTTAAGTATATTTATGCAAAAGTTTTCTGTGAGAAATGACAAAT			92290		
Qy	2218	gttaatatgcttataatgaataacagctcgtcttctccagagctccagagacatgtttaata			2277		
Db	92289	GTTAATATGCTTTATGAATTAACAGTGTCTTCCAGACTCCAGAGACATTTTATATA			92230		
Qy	2278	aagacaatgatcatgacccgaagaag			2301		
Db	92229	AAGCAATGAAATATATGACCGAAAG			92206		
LOCUS	AF326742	642 bp	mRNA	PII	28-FEB-2001		
DEFINITION	Cercopithecus aethiops sclerostin mRNA, complete cds.						
ACCESSION	AF326742						
VERSION	AF326742.1	GI:13161028					
KEYWORDS							
SOURCE							
ORGANISM							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
PUBMED							
REFERENCE							

[illegible]

RESULT	9
LOCUS	AF326741
DEFINITION	Rattus norvegicus sclerostin mRNA, complete cds.
ACCESSION	AF326741
VERSION	AF326741.1
KEYWORDS	GI:13161025
SOURCE	Norway rat. Rattus norvegicus Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathii; Muridae; Murinae;
ORGANISM	Rattus
REFERENCE	1 (bases 1 to 674) Brunkow,M.E., Gardner,J.C., Van Ness,J., Paepert,B.W., Kovacevich,B.R., Prohl,S., Skonier,J.E., Zhao,L., Sabo,P.J., Fu,Y.H., Alisch,R.S., Gillett,L., Colbert,T., Tacconi,P., Galas,D., Hammersma,H., Beighton,P. and Mulligan,J.T. Bone Dysplasia Sclerosteosis Results from Loss of the SOST Gene Product, a Novel Cystine Knot-Containing Protein Am. J. Hum. Genet. 68 (3), 577-589 (2001)
JOURNAL	11179006
PUBMED	2 (bases 1 to 674) Brunkow,M.E., Gardner,J.C., Van Ness,J., Paepert,B.W., Kovacevich,B.R., Prohl,S., Skonier,J.E., Zhao,L., Sabo,P.J., Fu,Y.H., Alisch,R.S., Gillett,L., Colbert,T., Tacconi,P., Galas,D., Hammersma,H., Beighton,P. and Mulligan,J.T. Direct Submission Submitted (07-DEC-2000) Genomics, Celtech Chiroscience Inc., 1631 220th St. SE, Bothell, WA 98021, USA
AUTHORS	Location/Qualifiers 1..674 /organism="Rattus norvegicus" /strain="Sprague-Dawley" /db_xref="taxon:10116"
TITLE	5'UTR CDS 33..674 /codon_start=1 /product="sclerostin" /protein_id="AAK13456.1" /db_xref="GI:13161026"
JOURNAL	/translation="MOLSAPCLACLIVHAFAVAVESOGMAFRNDATETIIPGIREYP EPQELNNOTNMRANGSRPHIRPPTKVSEYSCHELIITREYDGPSPRSKPYTE LVSGGGCPARLLPNALIGRWKMPNPDPICPDRIADRVOLDCGAAPSRRVR LVSCKRKRLTRFHNOSELKDGFERTARPDKRRPRPARAGANAQAENAY"
BASE COUNT	128 a 240 c 208 g 98 t
ORIGIN	
Query Match	23.3%; Score 536.2; DB: 94; Length 674;
Best Local Similarity	88.2%; Pred. No. 3.2e-93;
Matches	583; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
Oy	29 cccttccttgctgctgtacatgcagcctcccaatggcctgtctcgtctcgtctgctt 88
Db	14 CCTTCTCCCTTCGTGGCACATGCAGCACTCTCACTAGCCCCCTTGCGCTGCTTGT 73
Oy	89 acacacagccttcggtgaagtggaggcgagggatgtagcagcgcttaagaatgatgccac 148
Db	74 ACATCCACCCTTCGTTGCTGTGGAGAAGCCAGGGGGGCAACCTTCAGAAATGATGCCAC 133
Oy	149 ggaatcatcccacgagcctcgagagttaccoccgagcctcacacggagctcgagaaacaacaa 208
Db	134 AGAATATCATCCCGGAGACTCAGAGACTACCCAGAGCGCTCTCTCAGAGAACTAGAACAAACA 193
Oy	209 gacctgaacggcgagcgagaaacggagggcgagcctccacacacccccctltgagaccaaga 268
Db	194 GACCATGGAACCGGGCGGAGAACGGAGGAGAACCCCCCACCATCTTATGACACCAAAGA 253
Oy	269 cggtgcgaagtacaaactcgcggaactcgaacttcaacccgctacgaatgagacccaatggcgctg 328
Db	254 CGTGTCCGAGTAACACTGCTCCGCGAGCTGCACTACACCCGCTTCGTGACCGAGCGCCGTG 313
Oy	329 ccgcagcgcccaagcgcgttacaccgagctggtgtctcgcggcagatgcgagccggcgagcct 388

[illegible]

RESULT	10
AC073954	
LOCUS	
DEFINITION	AC073954 177744 bp DNA 17-NOV-2000
ACCSSION	Homo sapiens chromosome 17 clone RP11-147C10. WORKING DRAFT
VERSION	AC073954
KEYWORDS	SEQUENCE, 34 unordered pieces.
SOURCE	AC073954.3 GI:10988993
ORGANISM	HTG: HTGS_PHASE1; HTGS_DRAFT. human.
REFERENCE	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 177744) Waterston,R.H. The sequence of Homo sapiens clone Unpublished 2 (bases 1 to 177744) Waterston,R.H. Direct Submission Submitted (07-JUL-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA On Oct 25, 2000 this sequence version replaced gi:9186849.
JOURNAL	
TITLE	
AUTHORS	
REFERENCE	
JOURNAL	
COMMENT	

```

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: W05SC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H.NH0147C10
----- Summary Statistics -----
Sequencing vector: MJ3: 100%
Sequencing vector: plasmid: 0%
Chemistry: Dye-Primer ET7: 100% of reads
Chemistry: Dye-terminator Big Dye: 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 156769 bases at least Q40
Consensus quality: 163329 bases at least Q30
Consensus quality: 166780 bases at least Q20
Insert size: 185000; agarose-fp
Insert size: 183559; sum-of-contigs
Quality coverage: 3.29 in Q20 bases; agarose-fp
Quality coverage: 3.43 in Q20 bases; sum-of-contigs

```

\* NOTE: This is a 'working draft' sequence. It currently





Db 475 CTGATGACCTCTGTCAGTCAAGCGCTCACCGCTTCCAAACCAAGTCGGAGCTCAAG 534  
Qy 588 gactctggagccgagcgctcgccgaggaagcgccgagcgcccgcccgccg 647  
|||||  
Db 535 GACTTCGGCGGAGCGCGCGCGGAGAGGCTCCGAGACCGCGCGCGCGCGG 594  
Qy 648 agcgccaaagccagcgccgagctgagaaagcctaagag 691  
|||||  
Db 595 GGAGCCAAAGCAACGAGCGGAGCTGGAGAACGCTACTAGAG 638  
RESULT 12  
AF326738 532 bp mRNA MAM 28-FEB-2001  
LOCUS AF326738 Bos taurus sclerostin mRNA, partial cds.  
DEFINITION AF326738  
VERSION AF326738.1 GI:13161016  
KEYWORDS  
SOURCE COW.  
ORGANISM Bos taurus  
Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
Bovidae; Bovinae; Bos.  
REFERENCE 1 (bases 1 to 532)  
Brunkow,M.E., Gardner,J.C., Van Ness,J., Paepel,B.W.,  
Kovacevich,B.R., Prohl,S., Skonier,J.E., Zhao,L., Sabo,P.J.,  
Fu,Y.H., Alisch,R.S., Gillett,L., Colbert,T., Tacconi,P., Galas,D.,  
Hamersma,H., Beighton,P. and Mulligan,J.T.  
Bone Dysplasia Sclerosteosis Results from Loss of the SOS1 Gene  
Product, a Novel Cytosine Knot-Containing Protein  
Am. J. Hum. Genet. 68 (3), 577-589 (2001)  
11179006  
JOURNAL 2 (bases 1 to 532)  
PUBMED Brunkow,M.E., Gardner,J.C., Van Ness,J., Paepel,B.W.,  
REFERENCE Kovacevich,B.R., Prohl,S., Skonier,J.E., Zhao,L., Sabo,P.J.,  
AUTHORS Fu,Y.H., Alisch,R.S., Gillett,L., Colbert,T., Tacconi,P., Galas,D.,  
Hamersma,H., Beighton,P. and Mulligan,J.T.  
Direct Submission  
TITLE Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 1631  
JOURNAL 220th St. SE, Bothell, WA 98021, USA  
FEATURES  
source Location/Qualifiers  
1. 532  
/organism="Bos taurus"  
/db\_xref="taxon:9913"  
CDS  
/codon\_start=3  
/product="sclerostin"  
/protein\_id="AAK13453.1"  
/db\_xref="GI:13161017"  
/translation="MDAREIIPELGEYPEPLNKTNRARENGRPHHPRETKDA  
SEYSCRELHFRYTYDGCBSRAKRYTELVSGGCGPRKLLPNAIGKMKRPSGPRF  
CIPDRKARVOLLCPGGAARAKRVRKRVLVAACKCKRLTRFHNSLKDPEAPARPOT  
GKRLPRAGTKASNA"  
BASE COUNT 95 a 199 c 176 g 62 t  
ORIGIN

Query Match 20.3% Score 467.8 DB 7: Length 532;  
Best Local Similarity 93.5% Pred. No. 4.7e-80;

Matches 500; Conservative 0; Mismatches 32; Indels 3; Gaps 1;

Qy 136 agaatgctcgaagaatcattcccgagctcgagagatcccgagctccaccgagc 195  
|||||  
Db 1 AGAATGATCGCACAGAAATCATCTCCGAGCTGGCGAGTACCCGAGCTCTGCCAGAG 60  
Qy 196 tggagaaacaagaacatgaacccggcggaagagggcgcgctccaccaccct 255  
|||||  
Db 61 T--GAACAACAAGACCAATGACCGGCGGAGAACGAGGGAGACTCTCCACACCCCT 117  
Qy 256 ttgagcaaaagacgtctcgagatagcttcgagagctgactccaccgctacg 315  
|||||  
Db 118 TTGAGACCAAGAAGCGCTCCGAGTACAGCTGCCGGAGAGTGCACCTTACCCGCTACG 177

Qy 316 ccgatggccgtgctcgagagcgccaaagcagtcacagctgtgtgtctccgacagtgc 375  
|||||  
Db 178 CGATGGGCGCTGTGCGCGAGCGCCAGCGCGCTCACCGAGCTGATGCTCGGCGAGTCCG 237  
Qy 376 gcccgagcgctgtctgtcccaagcactgagcgcggaagtgtgtgtgtgtgtgtgt 435  
|||||  
Db 238 GCCCGCGCGCGCTGTGCGCGAGCGCCATCGCGCGCGCAAGTGTGTGTGTGTGTGTGT 297  
Qy 436 ccgacttcgctgtatcccgacccgctacccgagcgagcgctgtgtgtgtgtgtgtgt 495  
|||||  
Db 298 CCGACTTCGCTGTGATCCCGACCGCTACCGCGCGCGAGCGGTGTGTGTGTGTGTGTGT 357  
Qy 496 gt 555  
|||||  
Db 358 GCGGCG 417  
Qy 556 tcaccgcttcacaaacagtcagagctcaagagcttcggagcgagcgctgtgtgtgtgt 615  
|||||  
Db 418 TCACCTCGCTTCACACACCAAGCTCCAGAGCTTCGCGCGCGCGCGCGCGCGCGCGCG 477  
Qy 616 agaagggcgagagcgcgcgcccgcgcccgagcgagcgagcgagcgagcgagcgagcga 670  
|||||  
Db 478 AACCGCGCGGAGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 532  
RESULT 13  
AF326737 17423 bp DNA ROD 28-FEB-2001  
LOCUS AF326737  
DEFINITION Mus musculus sclerostin gene, complete cds.  
ACCESSION AF326737  
VERSION AF326737.1 GI:13161013  
KEYWORDS  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 17423)  
Brunkow,M.E., Gardner,J.C., Van Ness,J., Paepel,B.W.,  
Kovacevich,B.R., Prohl,S., Skonier,J.E., Zhao,L., Sabo,P.J.,  
Fu,Y.H., Alisch,R.S., Gillett,L., Colbert,T., Tacconi,P., Galas,D.,  
Hamersma,H., Beighton,P. and Mulligan,J.T.  
Bone Dysplasia Sclerosteosis Results from Loss of the SOS1 Gene  
Product, a Novel Cytosine Knot-Containing Protein  
Am. J. Hum. Genet. 68 (3), 577-589 (2001)  
11179006  
JOURNAL 2 (bases 1 to 17423)  
PUBMED Brunkow,M.E., Gardner,J.C., Van Ness,J., Paepel,B.W.,  
REFERENCE Kovacevich,B.R., Prohl,S., Skonier,J.E., Zhao,L., Sabo,P.J.,  
AUTHORS Fu,Y.H., Alisch,R.S., Gillett,L., Colbert,T., Tacconi,P., Galas,D.,  
Hamersma,H., Beighton,P. and Mulligan,J.T.  
Direct Submission  
TITLE Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 1631  
JOURNAL 220th St. SE, Bothell, WA 98021, USA  
FEATURES  
source Location/Qualifiers  
1. 17423  
/organism="Mus musculus"  
/strain="129Sv/J"  
/db\_xref="taxon:10090"  
/chromosome="11"  
/join(8036..8249,10742..11163)  
/product="sclerostin"  
/join(8036..8249,10742..11163)  
/codon\_start=1  
/product="sclerostin"  
/protein\_id="AAK13452.1"  
/db\_xref="GI:13161014"  
/translation="MPSLAPLCILYLHAARCAVEGGQMAFNDATVYIRIGEP  
EPPENQITNRNARENGRPHHPYPAKDVSEYSCRELHFRYTYDGCBSRAKRYTEL  
CSGCGPRRLPNAIGRVKRWMPNPGPRCTIDRTARVOLLCPGGAARVRVLY  
ASCKRLTRFHNSLKDPEETARPQKGRKPRGARGAANAAELENAY"  
BASE COUNT 4456 a 4306 c 4569 g 4092 t  
ORIGIN



```

Query Match      16.2%; Score 372.2; DB 94; Length 17423;
Best Local Similarity 85.9%; Pred. No. 7.2e-62;
Matches 413; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 233 aggggagctcccaacacccctttgagacaaagctgttcagatcagctcgcgga 292
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10707 AGGCTGACACCCCTTCACGATCCCTTCCTCCGAGATGTCCGAGTACAGTCCGCGA 10766

QY 293 gctgcaactcacccgctagctgacagatgagcgtgcccagcagcagcagatcagga 352
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10767 GCTGCACTACACCCGCTTCTCTGACAGACGCGCCATGCCAGCCGCAAGCCGGTACCGA 10826

QY 353 gctggtgtgtctccgcagctgagccgcgcgcgcgtctgtcccaagcagcagcgcg 412
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10827 GTTGTGTGTCTCCGGCAGTGGCCGCCCGCGCTGTCTGCCCAACGCCATGCGGCGCT 10886

QY 413 caagtggtgagcagctagtgaggccgagctcgtctgcatcccgacagcagcagcagca 472
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10887 GAAGTGTGTGGCCCGCAACGAGACCGGATTTCCGCTGCATCCGGATCCGTAACCGCGCA 10946

QY 473 ggcgtgtgagcagctgtgtgtcccggtgtgtgagcgccgcgcgcgcgcgcgcgcgcgc 532
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10947 GCGGCTGACAGCTGTGTGTGCGCCGGGGGCGCGCGCGCTGCGCAAGTGCCTGTGT 11006

QY 533 ggcgtgtgtgacagtgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 592
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 11007 GGCCTGTGTCAAGTGCAGCGCTCACCGCTTCCACAACTGTGAGAGCTTCAGAGACTT 11066

QY 593 cggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 652
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 11067 CGGCGCGAGACCGCGCGCGCGCGCGAGAAAGGTGTGCAAGCGCGCGCGCGCGGAGC 11126

QY 653 caaagcacaacccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 712
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 11127 CAAAGCAACCAAGCGAGCTGAGAAACGCTACTAGACGAGCGCGCGCTATGCAGCC 11186

QY 713 c 713
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 11187 C 11187

RESULT 14
AC068782_2
WPCOMMENT
Sequence split into 5 fragments LOCUS AC068782 Accession AC068782
Fragment Name Begin End
AC068782_0 1 110000
AC068782_1 100001 210000
AC068782_2 200001 310000
AC068782_3 300001 410000
AC068782_4 400001 494838
Continuation (3 of 5) of AC068782 from base 200001 (AC068782 Mus musculus chromosome 11)

```

```

Db 56735 GAAGTGTGTGGCGCCGCAACGAGACCGGATTTCCGCTGCATCCCGATGCTACCGCGCA 56794
QY 473 ggcgtgtgagcagctgtgtgtcccggtgtgtgagcgccgcgcgcgcgcgcgcgcgcgc 532
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 56795 GCGGCTGACAGCTGTGTGTGCGCCGGGGGCGCGCGCGCTGCGCAAGTGCCTGTGT 56854

QY 533 ggcgtgtgacagtgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 592
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 56855 GGCCTGTGTCAAGTGCAGCGCTCACCGCTTCCACAACTGTGAGAGCTTCAGAGACTT 56914

QY 593 cggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 652
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 56915 CGGCGCGAGACCGCGCGCGCGCGCAAGAGGTGTGCAAGCGCGCGCGCGCGGAGC 56974

QY 653 caaagcacaacccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 712
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 56975 CAAAGCAACCAAGCGAGCTGAGAAACGCTACTAGACGAGCGCGCGCTATGCAGCC 57034

QY 713 c 713
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 57035 C 57035

```

```

RESULT 15
LOCUS AC012296
DEFINITION Mus musculus chromosome 11 clone RP23-346P7 map 11, WORKING DRAFT
ACCESSION AC012296
VERSION AC012296.4 GI:8072530
KEYWORDS HTGS, PHASL1, HTGS_DRAFT.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine; Mus.
1 (bases 1 to 200727)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Mus musculus chromosome 11, clone RP23-346P7
Unpublished
2 (bases 1 to 200727)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boukhalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collamore,A.,
Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Haos,B., Headford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczy,J., Lien,C., Locke,K., MacDonald,P., Marquis,N.,
McEwan,P., McGuirk,A., McKernan,K., McLaughlin,J., Meldrum,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zodi,M.

```

```

TITLE
JOURNAL
COMMENT
Direct Submission
Submitted (22-OCT-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 25, 2000 this sequence version replaced g1:7329383.
All repeats were identified using RepeatMasker:
Smith,A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/MW/repeatmasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: 13205
Center clone name: 346_P7
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731

```

Consensus quality: 181484 bases at least Q40  
Consensus quality: 190159 bases at least Q30  
Consensus quality: 193745 bases at least Q20  
Insert size: 207000; agarose-fp  
Insert size: 197427; sum-of-contigs  
Quality coverage: 3.8 in Q20 bases; agarose-fp  
Quality coverage: 4.0 in Q20 bases; sum-of-contigs

-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 34 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1091: contig of 1091 bp in length  
1092 1191: gap of 100 bp  
1192 3065: contig of 1874 bp in length  
3066 3165: gap of 100 bp  
3166 4305: contig of 1140 bp in length  
4306 4405: gap of 100 bp  
4406 5983: contig of 1578 bp in length  
5984 6083: gap of 100 bp  
6084 8322: contig of 2239 bp in length  
8323 8422: gap of 100 bp  
8423 10098: contig of 1676 bp in length  
10099 10198: gap of 100 bp  
10199 12592: contig of 2394 bp in length  
12593 12692: gap of 100 bp  
12693 14684: contig of 1992 bp in length  
14685 14784: gap of 100 bp  
14785 17163: contig of 2379 bp in length  
17164 17263: gap of 100 bp  
17264 19900: contig of 2637 bp in length  
19901 20000: gap of 100 bp  
20001 23054: contig of 3054 bp in length  
23055 23154: gap of 100 bp  
23155 26011: contig of 2857 bp in length  
26012 26111: gap of 100 bp  
26112 29412: contig of 3301 bp in length  
29413 29512: gap of 100 bp  
29513 32991: contig of 3479 bp in length  
32992 33091: gap of 100 bp  
33092 36913: contig of 3822 bp in length  
36914 37013: gap of 100 bp  
37014 40880: contig of 3867 bp in length  
40881 40980: gap of 100 bp  
40981 45015: contig of 4035 bp in length  
45016 45115: gap of 100 bp  
45116 50321: contig of 5206 bp in length  
50322 50421: gap of 100 bp  
50422 54284: contig of 3863 bp in length  
54285 54384: gap of 100 bp  
54385 59430: contig of 5446 bp in length  
59431 59530: gap of 100 bp  
59531 64373: contig of 4843 bp in length  
64374 64473: gap of 100 bp  
64474 70503: contig of 6030 bp in length  
70504 70603: gap of 100 bp  
70604 77829: contig of 7226 bp in length  
77830 77929: gap of 100 bp  
77930 85976: contig of 8047 bp in length  
85977 86076: gap of 100 bp  
86077 91694: contig of 5618 bp in length  
91695 91794: gap of 100 bp  
91795 99208: contig of 7414 bp in length  
99209 99308: gap of 100 bp  
99309 107499: contig of 8191 bp in length  
107500 107599: gap of 100 bp  
107600 115731: contig of 8132 bp in length  
115732 115831: gap of 100 bp  
115832 123297: contig of 7466 bp in length

123298 123397: gap of 100 bp  
123398 133579: contig of 10582 bp in length  
133580 134079: gap of 100 bp  
134080 14631: contig of 10552 bp in length  
144632 144731: gap of 100 bp  
144732 159118: contig of 14387 bp in length  
159119 159218: gap of 100 bp  
159219 172437: contig of 13219 bp in length  
172438 172537: gap of 100 bp  
172538 200727: contig of 28190 bp in length.

Location/Qualifiers  
1. .200727  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/chromosome="11"  
/map="11"  
/clone="RP23-346P7"  
/clone\_lib="RPC1-23 Female Mouse BAC"  
1. .1091  
/note="assembly\_fragment"  
1192. 3065  
/note="assembly\_fragment"  
3166. 4305  
/note="assembly\_fragment"  
4406. 5983  
/note="assembly\_fragment"  
6084. 8322  
/note="assembly\_fragment"  
8423. 10098  
/note="assembly\_fragment"  
10199. 12592  
/note="assembly\_fragment"  
12693. 14684  
/note="assembly\_fragment"  
14785. 17163  
/note="assembly\_fragment"  
17264. 19900  
/note="assembly\_fragment"  
20001. 23054  
/note="assembly\_fragment"  
23155. 26011  
/note="assembly\_fragment"  
vector\_end:SP6  
26112. 29412  
/note="assembly\_fragment"  
29513. 32991  
/note="assembly\_fragment"  
33092. 36913  
/note="assembly\_fragment"  
37014. 40880  
/note="assembly\_fragment"  
40981. 45015  
/note="assembly\_fragment"  
45116. 50321  
/note="assembly\_fragment"  
50422. 54284  
/note="assembly\_fragment"  
54385. 59430  
/note="assembly\_fragment"  
59531. 64373  
/note="assembly\_fragment"  
64474. 70503  
/note="assembly\_fragment"  
70604. 77829  
/note="assembly\_fragment"  
77930. 85976  
/note="assembly\_fragment"  
86077. 91694  
/note="assembly\_fragment"  
91795. 99208  
/note="assembly\_fragment"  
99309. 107499

```

/note="assembly_fragment"
misc_feature      107600..115731
                  /note="assembly_fragment"
misc_feature      115832..115832
                  /note="assembly_fragment"
clone_end:r7

```

Query Match 16.2%: Score 372.2: DB 62: Length 200727;  
 Best Local Similarity 85.9%: Pred. No. 5.2e-62;  
 Matches 413; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

```

QY 233 agggcgagcctccaccaccccttgaagacaaagacgtgtccagaglacagtcgcgca 292
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 150286 AGGCTGCACAGCCCTCAGCATCCCTCTCCGAGATGTCGAGTAACGTCGCGCA 150345

QY 293 gctgcacctcaccgctacgtgacgagtgagcgctgacgacgacgacgacgacgac 352
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      150346 GCTGCACCTACACCGCTTCTGACAGACGCGCCATGCCAGCGCCAGCGTCACCGA 150405

QY 353 gctggtgtctccgagcagtgagcgccgagcgagcgtgtgtcccaagccatcgagcg 412
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 150406 GTTGGTGTCTCTCGGSCATGGCGCCCGCGGCTGCTGCCAAGCCATCGGCGCT 150465

QY 413 caagtggtgagcagctagtgagcgccgacttcgctgtcaltcccgacgctacgcgcga 472
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 150466 GAAGTGTGTGGCGCCGAGACGACCGGATTTCGCTGATCCCGATCGTACCGCGCA 150525

QY 473 ggcggtgagcagctgtgtgtcccggtgtgtgagcgccgagcgagcgagcgagcg 532
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 150526 GCGGGTGCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 150585

QY 533 ggcctgtgtcaagtgcaagcgctcaccgcttcacacacagtcgagctcaagagact 592
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 150586 GGCCTGTGTCAAGTGTCAAGCGCTTCCACCAACAGTGTGAGCTCAAGAGACTT 150645

QY 593 cgggacgagcgagcgctcgagcgaggaagcgagcgagcgagcgagcgagcgagcgc 652
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 150646 CGGGCGGAGACCGCGCGCGCGCGAGAGGGTGCAGAGCCGCGCGCGCGGGGAGC 150705

QY 653 caagcacaacagcgagctgaggaagcctactagagcccgccgagccctccacac 712
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 150706 CAAGCCACACAGCGGAGCTGAGAGACGCTACTAGAGCGAGCCCGCGCTATGCAGCC 150765

QY 713 c 713
      |
Db 150766 C 150766

```

Search completed: November 8, 2001, 23:49:26  
 Job time: 16136 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 8, 2001, 23:49:26 ; Search time 9526.57 Seconds  
(without alignments)  
3736.003 Million cell updates/sec

Title: US-09-668-021-5  
Perfect score: 2301  
Sequence: 1 agagcctgctactggaag.....caatgatcatgacgaaag 2301

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

rchcd: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenBmb1: \*  
1: gb\_da1: \*  
2: gb\_da2: \*  
3: gb\_da3: \*  
4: gb\_in1: \*  
5: gb\_in2: \*  
6: gb\_in3: \*  
7: gb\_om: \*  
8: gb\_ov: \*  
9: gb\_pat1: \*  
10: gb\_pat2: \*  
11: gb\_ph: \*  
12: gb\_pl1: \*  
13: gb\_pl2: \*  
14: gb\_pl3: \*  
15: gb\_pl4: \*  
16: em\_ba1: \*  
17: em\_ba2: \*  
18: em\_fun: \*  
19: em\_htgo\_hum: \*  
20: em\_htgo\_inv: \*  
21: em\_htgo\_rod: \*  
22: em\_htg\_hum1: \*  
23: em\_htg\_hum2: \*  
24: em\_htg\_hum3: \*  
25: em\_htg\_hum4: \*  
26: em\_htg\_hum5: \*  
27: em\_htg\_hum6: \*  
28: em\_htg\_hum7: \*  
29: em\_htg\_hum8: \*  
30: em\_htg\_inv1: \*  
31: em\_htg\_inv2: \*  
32: em\_htg\_other: \*  
33: em\_htg\_rod: \*  
34: em\_hum1: \*  
35: em\_hum2: \*  
36: em\_hum3: \*  
37: em\_hum4: \*  
38: em\_hum5: \*  
39: em\_hum6: \*  
40: em\_hum7: \*  
41: em\_in: \*  
42: em\_om: \*  
43: em\_or: \*

44: em\_ov: \*  
45: em\_pat: \*  
46: em\_ph: \*  
47: em\_pl: \*  
48: em\_rod: \*  
49: em\_sts: \*  
50: em\_sy: \*  
51: em\_un: \*  
52: em\_v1: \*  
53: gb\_sts1: \*  
54: gb\_sts2: \*  
55: gb\_sts3: \*  
56: gb\_sy: \*  
57: gb\_un: \*  
58: gb\_v1: \*  
59: gb\_v12: \*  
60: gb\_htg1: \*  
61: gb\_htg2: \*  
62: gb\_htg3: \*  
63: gb\_htg4: \*  
64: gb\_htg5: \*  
65: gb\_htg6: \*  
66: gb\_htg7: \*  
67: gb\_htg8: \*  
68: gb\_htg9: \*  
69: gb\_htg10: \*  
70: gb\_htg11: \*  
71: gb\_htg12: \*  
72: gb\_htg13: \*  
73: gb\_htg14: \*  
74: gb\_htg15: \*  
75: gb\_htg16: \*  
76: gb\_htg17: \*  
77: gb\_htg18: \*  
78: gb\_htg19: \*  
79: gb\_htg20: \*  
80: gb\_htg21: \*  
81: gb\_htg22: \*  
82: gb\_htg23: \*  
83: gb\_htg24: \*  
84: gb\_htg25: \*  
85: gb\_pr1: \*  
86: gb\_pr2: \*  
87: gb\_pr3: \*  
88: gb\_pr4: \*  
89: gb\_pr5: \*  
90: gb\_pr6: \*  
91: gb\_pr7: \*  
92: gb\_pr8: \*  
93: gb\_pr9: \*  
94: gb\_pr1: \*  
95: gb\_pr2: \*  
96: gb\_in4: \*  
97: gb\_pr10: \*  
98: em\_ba3: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2297.8	99.9	2323	89	AF326739 Homo sapi
2	2277.4	99.0	2329	9	AX056687 Sequence
3	2270	98.7	2296	89	AF331844 Homo sapi
4	2047.8	89.0	21501	89	AF326736 Homo sapi
5	2047.8	89.0	94752	85	AC003098 Homo sapi
6	2035.8	88.5	151780	72	AC055813 Homo sapi
7	1066.4	46.3	177744	75	AC073954 Homo sapi
8	610	26.5	642	89	AF326742 Cercopit

Query Match	99.9%	Score 2297.8;	DB 89;	Length 2323;
Best Local Similarity	99.98%	Pred. No. 0;		
Matches 2299;	Conservative	0; Mismatches	2; Indels	0; Gaps
1 agagcctgtgtcaacacgggaagtgtgcgttcacctctctgtgcgtgtacatgtacagctccac				
1 AGAGCCTGTGTCAACACGGGAAGTGTGCCTCTCTGTGCGTGTACATGTACAGCTCCAC				
61 tggccctgtgtcacaatctgtacctgtctgtgtacacaagcttcgltgtagtgtgaagggcag				
61 TGGCCTGTGTCAAAATGATGTCACAGGAATCATCCCGAGCTTCGTAAGTGTGAAGGCCAG				
121 ggtgtgcagagcgttcaagaatgtccacggaatcatctcgcgagcttcggaagtaaccg				
121 GGTGTGCAGAGCGTTCAAAGATGATGTCACAGGAATCATCCCGAGCTTCGTAAGTGTACCCG				
181 agcctaccggaagctgtgaagaaacaagacatgtaacccggcggaagaaaggaagggcg				
181 AGCCTCACCGGAGCTGTGAAGAACAAAGACCATGTAAACCGGCGGAGAACGAGGCGGC				

QY	301	tcaccgagctacgtgacagatggccgctgtccgcaagcgcaagccggtcaaccgagctgtgtc	360
Db	301	TCACCCGCTACGAGACCGATGGGCGCTGCCACGCGCAAGCGGATCACCGACTGTGT	360
QY	361	gtcccgagccaatgtcgagcccgcgcgctgtcgtcccaacgccaatcgtgcgcggaagtgtt	420
Db	361	GTCTCCGCGCAAGTCCGCGCCCGCGCCCGCTGCTCCACAGCCATCGCGCGCAAGTGT	420
QY	421	ggtgacctagtgtgagccgcaattccgctgtcatctcccgacgactaacgcgacgcaagctgtac	480
Db	421	GCGGACCTAGTGGGCCCGACTTCGCGCTGCAATCCCGAACGCTAACGCGCGCCACGCGTGC	480
QY	481	agctctgtgtctcccggtgtgtgtagagcgccgacgacgcaagttgcctgtgtgtcctgtc	540
Db	481	AGCTGCTGTGTCTCCGGTGTGTGAGGCGCGCGCGCCAAAGTGTGCGCTGTGTGCTCTGCT	540
QY	541	gcaaatgtgcaagcgctctcacccgcttcccaacccagtgtgagcttcaaggaatttgggaacg	600
Db	541	GCAAGTGCAAAGCGCTCTCACCCGCTTTCACAACCAAGTGTGGAGCTTCAAGGAACTTGGGACCG	600
QY	601	agacgcgctcgccgacgaagaagtcggaagccgagcccgccgacgacgcgagtcgccaagaagcca	660
Db	601	AGGCGCGTCCGCCGCGCAAGAGGCGCGGAAGCGCGGCCCGCGCCCGAGCGCCAAAGCCA	660
QY	661	accgaagccgaagctgtgaaagcctactagagcccgcccgcgccctccccaaccgacgagc	720
Db	661	ACCAGCGCGAGCTGTGAGAACGCTACTAGAACCCGCGCCGCGCCCTCCACACGCGGCGGCG	720
QY	721	gccccgagccctgaaccacgagccccaacatttctgtcctctcgagcggtgtgtatgttttat	780
Db	721	GCCCCGCGCCCTGAACCCGCGCCCAATTTCTGTCTCTGTGGCGTGGATTGATTTGTTTAT	780

QY	1861	tcacacatgataatgacatactcttaacataaagaattatctggggaataacataaagt	1920
Db	1861	TCCTNACTGATGATGGCATATCTTACACTAAAGAAATTTATTTGGGGGAAAAACATCAAGT	1920
QY	1921	gcgtacacataatgcggaataactgcagagacataaagtctgcacacccaaaaactctttgaa	1980
Db	1921	GCCTGACATATGCTGTAACAACTGCAAGCATTAATAGCTGCCACCCAAAATCTTTTGGAA	1980
QY	1981	aatcatcttcacagacaccccttaacttctctgtgtagcttlttaattgtttaaaaaaaagt	2040
Db	1981	AATCATCTTTCACACAAACCTCTTCTTCGTGTGATGTTTAAATGTATAAAAAAAAAGT	2040
QY	2041	tttaaacagaagacacatagacataatgaagaagctgcagagaccggtcgttctttctggcaatc	2100
Db	2041	TTTAAACAAAGACACATGACATATGAAACCCGCAAGACGTCGTCTTTTGGCAATTC	2100
QY	2101	ttcacagctggagactctgcacagaagaatgaatagtgattctttaagaagcttaagttcat	2160
Db	2101	TTCCACGTGGGACTGTGCCAAGATGAATGAAGTGTGTTTAAAGAGTTAATTTATCAT	2160
QY	2161	attcatcttctcaactaagtattatctgtgaagaagttcttctctgtatagaatagacaatgct	2220
Db	2161	ATTATATTTTCTCAGCTATAGTTATTTATGCAAAAGTTTCTTGTAACAGATGACAAATCT	2220
QY	2221	aataatgcttataaataaagaatcgtgtcttcacagatccagaagactttataaga	2280
Db	2221	AATATGCTTTATGATTAACAGTCTGTCTTCCAGATCCAGACGACATGTTATTAATAG	2280
QY	2281	acaaatgaatcatgaccgaag 2301	
Db	2281	ACAATGAATCATGACCGAAG 2301	
RESULT	2		
AX056687	2329 bp	DNA	PAT 17-JAN-2001
LOCUS	AX056687		
DEFINITION	Sequence 19 from Patent WO0075317.		
ACCESSION	AX056687		
VERSION	AX056687.1	GI:12309667	
KEYWORDS	.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	1. (bases 1 to 2329)		
AUTHORS	Boistein,D.A., Goddard,A., Gurney,A.L., Smith,V., Matanabe,C.K. and		
TITLE	Compositions and methods for the treatment of tumor		
JOURNAL	Patent: WO 0075317-A 19 14 DEC 2000.		
FEATURES	Genentech, Inc. (US)		
source	Location/Qualifiers		
	1..2329		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
BASE COUNT	587 a 645 c 612 g 485 t		
ORIGIN			
Query Match	99.0%;	Score 2277.4;	DB 9; Length 2329;
Best Local Similarity	99.7%;	Pred. No. 0;	
Matches 2281;	Conservative 0;	Mismatches 6;	Indels 0; Gaps 0;
QY	14	cttgaagatgagctgtgcctctctctctctgtgtacatagacatccacatgcctgtgct	73
Db	28	CTTGAAGATGAGCTGCTCCTCTCTCTCTGTGTACATGACATCCACATCGCCCTGTGCT	87
QY	74	catctgctgcgtgtacacagaaccttcgcgtgtatgtgagagccagggatgacagcgct	133
Db	88	CATCTGCTGCCTGTGACACAGCTTCCTGCTATGAGAGGCGCAAGGGATGACAGCGCT	147
QY	134	caagaatgctgcaagaatcatctccgagactcggaagtaacccgagactccacccga	193





RESULT	3
LOCUS	AF331844
DEFINITION	AF331844 Homo sapiens SOST (SOST) mRNA, complete cds.
VERSION	AF331844
KEYWORDS	AF331844.1 GI:13236417
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 2296) Bailemans,W., Ebeling,M., Patel,N., Van Hul,E., Olson,P., Dioszegi,M., Laczka,C., Wuyts,W., van den Ende,J., Willemis,P., Pas-Alves,A.F., Hill,S., Bueno,M., Ramos,F.J., Tacconi,P., Dijkers,F.G., Stratakis,C., Lindpaintner,K., Vickery,B., Foerzler,D. and Van Hul,W. Increased bone density in sclerosteosis is due to the deficiency of a novel secreted protein (SOST) Hum. Mol. Genet. 10 (5), 537-543 (2001)
TITLE	2 (bases 1 to 2296) Bailemans,W., Ebelling,M., Patel,N., Vickery,B., Foerzler,D. and Van Hul,W.
JOURNAL	Direct Submission Submitted (22-DEC-2000) Medical Genetics, University of Antwerp, Universiteitplein 1, Antwerp B2610, Belgium
PUBLISHED	Location/Oallifiers 1. .2296 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="17" /map="17q12-q11" 1. .2296 /gene="SOST" 38. .679 /gene="SOST" /note="mutations may cause sclerosteosis" /codon_start=1 /product="SOST" /protein_id="AAK16158.1" /db_xref="GI:13236418" /translation="MOLPLAICIVLHYFAFYVDSQGNQAFKNDAETILPEIGEP EPPLEENKTNNNAENGGRPRPHHPETFKDYSEKSCHLEHTFRVTYTPGCRSAKYVE LVASGGCCPARLLPNAIGRGWMPSPDFRCIDRTIADGVOLLCGEGAPRRKKRV LVAACKCKRLRFRFNOSLEKDFGEARPOKRPKRPPARSAKANQALEWAY"
FEATURES	source
GENE COUNT	576 a 631 c 607 g 482 t
GC IN	
Query Match	98.7%; Score 2270; DB 89; Length 2296;
Best Local Similarity	99.7%; Pred. No. 0;
Matches 2284; Conservative	0; Mismatches 5; Indels 1; Gaps 1;
OY	11 ctactctgaagtggtgcgtgctcctcctcttgcgtgtatcaacatgcattccaccatggccctgtg 70
Db	1 CTACTGGAAGGTGGCGTGCCTCTCTTGCTGTACCATGTACGATGCCACTGGCCTGTG 60
OY	71 tctctctgcttcctgcgtgtacacacagaccttcgttgtatgtagaaggcgaggtgtgcaagc 130
Db	61 TCTCTGTCCTCCTGCTGTGTAACAACAAGCCTTCCGTGTATGTGGAGGGCCAGAAGGTGGCAAGC 120
OY	131 gtccaagaatgatgccacgcgaaatcatccgcgagtctcgagaagtacccccgagctctcaac 190
Db	121 GTTCAAAGATGTATGCCACGCAAGTAATCATTCGCCAGCTCGAGTAGTACCCTCCACAC 180
OY	191 ggagctgtagaacacaagaacatatgaacggcgaggaagaacggaagggcgctcccacaca 250
Db	181 GGAGCTGGGAACACMACAGACCATATAACCGGCGGAGAACGAGGCGGCGCTCCCCACCA 240
OY	251 ccctcttgagaccaagagctgttccgaagtacagctgtccgagctgtgacttaaccgcta 310
Db	241 CCCCTTTGAGACCAAGAGCTGTCCGAGTAGACGTCCGCCCGAGCTGCACATTACCCGGTA 300

[illegible]

OY	1391	tcacgaatc	taagaggttg	tatgtatgtctcttc	tcagacagccaaga	tgaaacaacaacga	1450	
OY	1391	tcacgaatc	taagaggttg	tatgtatgtctcttc	tcagacagccaaga	tgaaacaacaacga	1450	
Db	1381	TCcGAAATTCAGAGCTTGTATGCTCTCTTC	GTcGACcCCAAAGATAAAAACAAACAGAA				1440	
OY	1451	aaaaaa	taagaagtlcat	ttatgtgc	tcacatat	ttacgcgtcagcaaaactc	tcygaaga	1510
Db	1441	AAAAAAAGTAAAGAGTCTATTATTGTGCTGACATATTATTAGCGCTGACAAACTCTGGAGA					1500	
OY	1511	agcctatgcctctccagagccttgctctcccgabctt	cttggtcactctcagccctccactct				1570	
Db	1501	AGcTATcTcTCTCCACAGcCTGcCTTCcCGGATGTTbGCTACCTCCACACcCTTCATCT					1560	
OY	1571	caagaaga	taaca	tcata	ctcttg	gggtaagaaagagagtg	tcagaggttg	1630
Db	1561	CAAGAAGATTAACATCATTCATCTGGGGTAGAAAAAGAGAGCTCCGAGGGTGGTGGAGGG					1620	
OY	1631	atagaat	taacatccgccc	ccaactctcccaagaagacga	tcctctcccccagccatg	cc	1690	
Db	1621	ATTAAGATTAACATCTCCGCCCCCAACTCTCCAAAGAGCGATGCTCTCCCGAGCCATPAACC					1680	
OY	1691	atgctttaa	agcaactctccgaagaag	tgaaaga	ttcaagaacactc	ggccttcagc	1750	
Db	1681	ATGTTTTAAAGTACCTTCGAGAGAGAATGAAGTTCAAGACACTGGCCTTSCAGGC					1740	
OY	1751	ccgagga	gcagccatca	caaaactca	cagacga	ccagacatccctlttg	agaacgcgttc	1810
Db	1741	CCGAGGAGAGAGCCATCATCAAACTCAACAGACCGACATCCCTTTGAGACACCGCCCTTC					1800	
OY	1811	tgccacac	actcaea	ggagacatctctgc	ctcagaanaacagctctc	taagctcttcacatgt	1870	
Db	1801	TGCCACCACTCAACGAGACATATTCTGCGCTAGAGAAACACTCTTACAGCTCTTACATGT					1860	
OY	1871	gatygc	atalc	ctac	actaaaga	tatatat	tg	1930
Db	1861	GATGGCATATCTTACACTAAAGAAATATATTTGGGGGAAAACTCAAGTGTGTACATA					1920	
OY	1931	tgctgaga	aacttg	agagca	taata	gtctgc	ccaccaaaactctltg	1990
Db	1921	TGCTGAGAACTCGAGAGCATATA-CTGCGACCCAAAATCTTTTGAAATCATTTCC					1979	
OY	1991	agacaac	ctct	actctc	tg	tgtagttt	taattg	2050
Db	1980	AGACAACCTCTTACTTCTGTGTAGTTTTTAATTTGTTAAAAAAAAGTTTTAAACAGA					2039	
OY	2051	agacat	gacat	atagaag	ccctgc	cagagac	tg	2110
Db	2040	AGCCATATGACATATGAAAGCTCGAGAGCTGCTGTTTTTTTGGCAATTTCTCCAGCTGG					2099	
OY	2111	gact	tg	tccaca	aga	atg	aa	2170
Db	2100	GACTTGTCACAAAGATGAAGTAGTGTTTTTAAAGAGTAAAGTTACTATTATTATTTC					2159	
OY	2171	tcact	taagt	atattat	gacaaag	tttctc	tg	2230
Db	2160	TCACCTTAAGTATTATATGCAAAAAGTTTTCTTCTTAAGAGATACAAATGTATATTGCTT					2219	
OY	2231	tatgaat	taac	agctc	gttcttc	ccagagtc	cagagacat	2290
Db	2220	TATGAATTAACAGCTGTCTCTCCAGAGTCACAGACATTGTTAATAAAGCAATGATTC					2279	
OY	2291	atgac	g	a	a	a	a	2300
Db	2280	ATGAATAAA 2289						

RESULT	4			
AF326736	AF326736	21501 bp	DNA	28-FEB-2001
LOCUS	Homo sapiens	sclerostin gene,	complete cds.	
DEFINITION	AF326736			
ACCESSION	AF326736.1	GI:13161010		
VERSION				
KEYWORDS	human.			
SOURCE				

	ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	TITLE	1 (bases 1 to 21501) Brunkow,M.E., Gardner,J.C., Van Ness,J., Paepert,B.W., Kovacevich,B.R., Prolli,S., Skonler,J.E., Zhao,L., Sabo,P.J., Fu,Y.H., Alisch,R.S., Gillett,I., Colbert,T., Tacconl,P., Galas,D., Hamernsa,H., Beighion.P. and Mulligan,J.T. Bone Dysplasia Sclerosteosis Results from Loss of the SOSF Gene Product, a Novel Cystine Knot-Containing Protein <i>Am. J. Hum. Genet.</i> 68 (3), 577-589 (2001)		
JOURNAL	PUBMED	11179006		
REFERENCE	AUTHORS	2 (bases 1 to 21501) Brunkow,M.E., Gardner,J.C., Van Ness,J., Paepert,B.W., Kovacevich,B.R., Prolli,S., Skonler,J.E., Zhao,L., Sabo,P.J., Fu,Y.H., Alisch,R.S., Gillett,I., Colbert,T., Tacconl,P., Galas,D., Hamernsa,H., Beighion.P. and Mulligan,J.T. Direct Submission Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 1631 220th St. SE, Bothell, WA 98021, USA		
FEATURES	source	Location/Qualifiers 1..21501 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="17" /map="17q21" join<11961..12180,14939.>.15360) /product="sclerostin" join(11961..12180,14939..15360) /codon_start=1 /product="sclerostin" /protein_id="AAK13451.1" /db_xref="GI:13161011" /translation="MQPLALCLVLIHFARFYVGQMAFKNDATLILPELGEP BPPLLENKKTMNAENGSRPPIHPETKDSEYSESLRHFRITYTDPSCAKPYIE LVSSQCQPARLLPNALRGKWMPSPGPFRCIPDRIROROLCPGEGLPRARKVR IVASCKRRLRFRIINOSLEIKDFTGARPOGRKRPPRARSAKANOAELBNAY"		
BASE COUNT	5546 a	5259 c	5419 g	5277 t
ORIGIN				
Query Match	Best Local Similarity	89.0%;	Score 2047.8;	DB 89; Length 21501;
Matches 2061;	Conservative	0;	Mismatches 22;	Indels 0;
Gaps 0;				
Db 14890	CAGGGGTGCACAGCGCGGCCACCCTCAGCCGGCTCTCTCACAGAAGCGTCCGAG	14949		
Qy 219	cgggaggaagaacaggagcgccctcccaccacctttagaaccaaaagcagtgcag	278		
Db 14890	CAGGGGTGCACAGCGCGGCCACCCTCAGCCGGCTCTCTCACAGAAGCGTCCGAG	14949		
Qy 279	tacaagtcgccagatctgaattcaaccgcctaagtgaacgatggcgtagcgagcgcc	338		
Db 14950	TACAGCTGCCGCGAGCTGCATTCAACCGGTACGTCAGCATGGGCGCTGCCACGCGC	15009		
Qy 339	aagccggtcaccaaagttgtgtgtccggacagtgagccccggcgccgtctgtgccaac	398		
Db 15010	AAGCGGTCAACCAAGTGTTGCTGTCTCGCGCANATGGCGCGCGCTGTGCCCCAAC	15069		
Qy 399	ggcatcgcccgcgagtaagtgtgtgcacactagtggcccgacttcogctgtacatcccgac	458		
Db 15070	GCCATCGCGCGCGCAAGTGTTGCGACACTTGTGGCGCCGACTTCGCTTCCATCCCAGAC	15129		
Qy 459	cgcctacgcgcgacgagcggtgaagctgtgtgtcccgagtgatgaagcgccgcgcgcac	518		
Db 15130	CGCTACCGCGCGAAGCGCTGCACTGCTGTGTCCCGTGTGAAGCGCGCGCGCGCC	15189		
Qy 519	aaggtgcgccttgtgtgtctcgtgtcaaatgtcaacgcgcactcaaccgcgtctccaacaaagtcg	578		
Db 15190	AAGGTGCGCTGTGCTCTGTGTAAGTCAAAGCGCGCTTCCAACAACCACTGCG	15249		
Qy 579	gaagcctaagaagacttcggaacggaagcgccgttcgacgaagaagcgccgaagcgccgcgc	638		
Db 15250	GAGCTCAAGACTTCGGGACCGAGGCGCGCTGCGCGCAGAAGGGCGCGAAGCGCGCGCC	15309		

[illegible]

Db	Accession	Sequence	Length	Position
Db	1630	GTGAAGGTTCAAGGACACTGGCTTGCAGAGCCGAGGAGACACCATTCCAAACTCACA	16449	
Oy	1779	gaaccagacatcccttttgagaaacccgctctcgcacacatcaaggacatcttcgc	1838	
Db	16450	GACGAGCACATCCCTTTTGAGACACCGCCCTCCGCCACCCATCCAGGACACATTTCTCG	16509	
Oy	1839	ctagaacacgctcttctactgctccttcaatgtagacgacatctcttaacaaagaatat	1898	
Db	16510	CTAAGAAACAGCTTTCTTACGCTCTTACCATGATGATGGCATATCTTTACCTAAGAAATAT	16569	
Oy	1899	tattgggggaaaaaacacacaaagtcgtcgtcatatgcttgagaaacgagacataatagct	1958	
Db	16570	TATTGGGGGAAAACACTACAGTGTCTTACATATAGCTTGAAACAGCAGACATTAATAGCT	16629	
Oy	1959	gcaccccaaaatccttttgaaatcaatctccagacaacotcttactctctgtagatt	2018	
Db	16630	GCCACCCAAAATCTTTTGAAAATCATATTTCCAGACAACCTCTTACTTCTGTGTAGTTT	16689	
Oy	2039	ttaactgttaaaaaaaaagttttaaagaagacacatgacatataagaacccgcagga	2078	
Db	16690	TTAATGTGTAAAAAAAAGTTTAAACAGAGACATGACATATGAAAGCCTGCAGGA	16749	
Oy	2079	ctggtcgttttttttggaatctctccaaagtgagcttgccaaagaatgaagtatg	2138	
Db	16750	CTGGTCCTTTTGTGGCAATCTCCACCTGGAGCTTCTCCACAAGATGAATGATGTGG	16809	
Oy	2139	ttttaaagagttaaagttacatatctatcttctccacttaagttatcttaagaaatctt	2198	
Db	16810	TTTTTAAGAGTTAAGTTACATATTTATTTTTCCTTAAGTTATTTATCCAAAAGTTT	16869	
Oy	2199	tcttgtagaagtagacaaagtttaattcttgctttagaattaaagcttgcttctccaag	2258	
Db	16870	TCTTGTGACAAATGACAAATGTTAATTTCTTTATGATTTAAACGTCTGTCTTCCAGAG	16929	
Oy	2259	tccaagacatctgttaataagaacaatgaatcagaccgaag	2301	
Db	16930	TCCAGACACATTTGTTATTAAGACATGATCATATGACCCGAAG	16972	
RESULT 5	AC003098/c	94752 bp	DNA	29-JAN-1998
LOCUS	AC003098	Homo sapiens chromosome 17, clone HRPc905M1, complete sequence.		
DEFINITION	AC003098			
ACCESSION	AC003098.1	GI:2822155		
VERSION				
KEYWORDS				
SOURCE		human.		
ORGANISM		Homo sapiens		
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
REFERENCE		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS		1 (bases 1 to 94752)		
TITLE		Birren,B., Fasman,K., McKernan,K., Nusbaum,C., Richardson,P. and		
JOURNAL		Homo sapiens chromosome 17, clone HRPc905M1		
REFERENCE		Unpublished		
AUTHORS		2 (bases 1 to 94752)		
TITLE		Birren,B., Fasman,K., McKernan,K., Munro,C., Nusbaum,C.,		
JOURNAL		Richardson,P., Lander,E., Balow,N., Barran,N., Cantu,C., Chang,A.,		
REFERENCE		Cooke,P., Daly,M.J., Devon,K., Dewar,K., Duketle,B., Forrest,C.,		
AUTHORS		Galpe,D., Genshlimer,S., Geradley,K., Gilmarlin,T., Hagos,B.,		
TITLE		Halden,I., Harris,K., Howland,J.C., Huang,J., Hui,L., Jacobot,L.,		
JOURNAL		Kilby,A., Lane,M., Mackenzie,J., Marquis,N., Miller,M., Jacotot,L.,		
REFERENCE		Molla,M., Morrow,J., Nachman,A., Naylor,J., O'Connor,T., Olotu,A.,		
AUTHORS		Peterson,K., Roberts,D., Rollins,G., Sarnaik,A., Shu,P., Shyam,R.,		
TITLE		Silivelli,J., Stone,C., Strickland,C., Sydney,K., Tang,L.,		
JOURNAL		Zemteva,I. and Zody,M.		
REFERENCE		Direct Submission		
AUTHORS		Submitted (10-NOV-1997) Whitehead Institute/MIT Center for Genome		
TITLE		Research, 320 Charles Street, Cambridge, MA 02141, USA		
JOURNAL		3 (bases 1 to 94752)		
REFERENCE		Birren,B., Fasman,K., McKernan,K., Nusbaum,C., Richardson,P.,		
AUTHORS		Lander,E., Allen,N., Baker,J., Baldwin,J., Barran,N., Beckery,R.,		

Boutwell, C., Byrne, S., Cantu, C., Castle, A., Cooke, P., Daly, M. J.,  
 Depierre, E., Devon, K., Dewar, K., Durelle, B., Etemadi, S.,  
 Ferreira, P., Forrest, C., Gage, D., Gardyna, S., Gensheimer, S.,  
 Gierdycz, K., Gilmartin, T., Gray, D., Hagos, B., Harris, K.,  
 Horton, L., Howland, J. C., Hui, L., Jacotot, L., Litton, L.,  
 Mackenzie, J., Marguis, N., McKean, P., McGurk, A., Meldrum, J.,  
 Molla, M., Morris, W., Morrow, J., Nachman, A., Naylor, J., O'Connor, T.,  
 Pavlin, B., Peterson, K., Ranganath, S., Riley, R., Roberts, D.,  
 Rollins, G., Rossello, R., Roy, A., Shyam, R., Soohoo, S., Stillwell, J.,  
 Stone, C., Strickland, C., Sydney, K., Tang, L., Vassiliev, H., Vo, A.,  
 Wagner, A., Wheeler, J., Wu, Y., Ye, W. J., Zemtseva, I., Zhao, J. and  
 Zody, M.

TITLE Direct Submission  
 JOURNAL Submitted (29-JAN-1998) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT On Jan 29, 1998 this sequence version replaced gi:2665504.  
 The Staden databases, finishing information, and all  
 chromatographic files used in the assembly of this clone are  
 available from our anonymous ftp site.

All repeats were identified using RepeatMasker: Smit, A.F.A. &  
 Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html.

Location/Qualifiers

1.94752

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="HRC905N1"

/clone\_1lb="Peter de Jong/ human PAC library"

/map="17"

/chromosome="17"

86. .439

/rpt\_family="LINE2"

complement(972. .1262)

/rpt\_family="AluSx"

complement(1884. .1912)

/rpt\_family="(CA)n"

1912. .1959

/rpt\_family="(GA)n"

2633. .2681

/rpt\_family="MIR"

2701. .3002

/rpt\_family="AluSx"

complement(3874. .4008)

/rpt\_family="LINE2"

4513. .4543

/rpt\_family="AT-rich"

complement(5202. .5237)

/rpt\_family="(GA)n"

complement(5310. .5406)

/rpt\_family="MIR"

complement(5829. .5868)

/rpt\_family="GC-rich"

6485. .6592

/rpt\_family="MIR"

complement(6973. .7029)

/rpt\_family="MIR"

complement(10402. .10426)

/rpt\_family="GC-rich"

complement(11421. .11496)

/rpt\_family="LINE2"

complement(11505. .11804)

/rpt\_family="AluSx"

11811. .11879

/rpt\_family="AT-rich"

11880. .12093

/rpt\_family="Aluub"

complement(12104. .12476)

/rpt\_family="LINE2"

12478. .12773

/rpt\_family="AluSx"

complement(12775. .12917)

/rpt\_family="LINE2"

repeat\_region

complement(13088. .13510)

/rpt\_family="MLT1C"

repeat\_region

complement(13544. .13847)

/rpt\_family="AluSx"

repeat\_region

complement(14110. .14776)

/rpt\_family="LINE1"

repeat\_region

complement(14777. .15077)

/rpt\_family="AluSx"

repeat\_region

complement(15079. .15282)

/rpt\_family="LINE1"

repeat\_region

15954. .16024

/rpt\_family="AT-rich"

repeat\_region

18031. .18178

/rpt\_family="MIR"

repeat\_region

20398. .20769

/rpt\_family="LINE2"

repeat\_region

20893. .21190

/rpt\_family="AluSg"

repeat\_region

complement(21275. .21473)

/rpt\_family="AluSg"

repeat\_region

complement(21485. .21556)

/rpt\_family="(CA)n"

repeat\_region

21567. .21906

/rpt\_family="MER2"

repeat\_region

21907. .21952

/rpt\_family="MLT2CB"

repeat\_region

21974. .22115

/rpt\_family="MIR"

repeat\_region

22238. .22540

/rpt\_family="AluSx"

repeat\_region

22577. .22633

/rpt\_family="LINE2"

repeat\_region

complement(23190. .23270)

/rpt\_family="LINE2"

repeat\_region

24196. .24324

/rpt\_family="FLAM\_A"

repeat\_region

24329. .24643

/rpt\_family="AluSp"

repeat\_region

25542. .25698

/rpt\_family="LINE2"

repeat\_region

25805. .26001

/rpt\_family="MIR"

repeat\_region

26060. .26174

/rpt\_family="LINE2"

repeat\_region

26343. .26649

/rpt\_family="AluSx"

repeat\_region

complement(27491. .27783)

/rpt\_family="AluSg"

repeat\_region

27862. .27906

/rpt\_family="LINE2"

repeat\_region

28068. .28114

/rpt\_family="MIR"

repeat\_region

complement(28329. .28360)

/rpt\_family="AT-rich"

28371. .28652

/rpt\_family="AluSg"

complement(29498. .29604)

/rpt\_family="GC-rich"

30074. .30181

/rpt\_family="MIR"

complement(30181. .30249)

/rpt\_family="LINE2"

30444. .30517

/rpt\_family="MIR"

complement(31852. .32135)

/rpt\_family="AluSg"

complement(32675. .33003)

/rpt\_family="MLT1A1"

complement(33042. .33082)

/rpt\_family="LINE2"

33363. .33459

/rpt\_family="MER5B"

complement(35475. .35524)

repeat\_region

[illegible]

Db	5680	TTGAGGGGTCACACGGGACGAGGGGAATTGAGAGTCACAGACACTGGACACGACCC	5621
Qy	939	ccgcctctggggccgcctacctgtctgcgtcccaacttcagagagcaaaatgaaagca	998
Db	5620	CCGCGCTCTGGGGCCGCCCTACCTTTGCTGTCGCCACTTCAGAGAGGACACAAATGGAACGA	5561
Qy	999	ttttcaaccgcttggggtttttaaggagcgggtgttggagatgggaaagttccaggagctgct	1058
Db	5560	TTTTTACCGCCCTGGGGTTTTTAAGGGACCGGTGTGGAGTGGGAAAGTCCAGGGACTCGT	5501
Qy	1059	taagaagaatttgataagaattcccccttgacacccgcgtcccatcaagaagccctggagggct	1118
Db	5500	TTAGAAATTTGGATTAGATTTCCTCCCTTGCACTCGCTGCCCATCAGAAAGCCTTGAGGGCT	5441
Qy	1119	gcccgagacacaagaacttgggggccaactgtatagatgtgtttctaagttccctgtcgtccact	1178
Db	5440	GCCCGACACACAAAGCTGGGGCAACTGTAGATGTGGTTTACGTCCTGGCTCGCACT	5381
Qy	1179	aactgtcgtgttaaccttgaacttaacaactctcccttggggccctcaattccactttgt	1238
Db	5380	AACCTTGCTGTGTAACCTTGAACTACACAAATTCCTTCGGGACCTCAATTTCCACTTGT	5321
Qy	1239	aaaatgaaggttggaggtgtgggaataagaaatctccagagagacatttgcataatgaattccaaag	1298
Db	5320	AAAATGAGGGGTGAGAGGTGGGAATAGATCTCGAGAGAGACTATTGGCATATGATTTCCAAAG	5261
Qy	1299	actccagtgccttttgaatgtggcagaggttgagagagagagagagaaagagagaaatgaa	1358
Db	5260	ACTCCAGGCGCTTTTGATATGGCGAGAGCTAGAGAGAGAGAGAAAGAGACAGATGAA	5201
Qy	1359	tgcacttgcattgaattcaagttccaaagttcaatccagaattcaagatttgaatgtccct	1418
Db	5200	TGCATTTGCATTGATTTCACTAGTCCCAAGTCACTTCAGAAATTCAGAGTTTGATGCTCTCT	5141
Qy	1419	tctgcagccaaagaatgaaaaaacaagaaaaaaaatgaagattcatalttaatgtct	1478
Db	5140	TCTGCACGCCCAAAATGATGAAAAACAAACGAAAAAAAAGTAAAGTCTATTATTGTGCT	5081
Qy	1479	gacattattacggcttgcgaacactcccttggaagaagctatagtcgtctccagccctgtccctc	1538
Db	5080	GACATATTTACGGCTGACAAACCTCTTGAGAAAGCTATGCTGCTTCCACGCTGGCTTCC	5021
Qy	1539	ccggaattgttgctacctccaccctccatctcaagaataataacatcatcaatcttggggtta	1598
Db	5020	CCGGATGTTTGGCTACCTCCACCCCTTCATCTCAAGAAATTAACATCATCATCTTGGGGTA	4961
Qy	1599	gaaaggggaaggtgcgaaggttgttgggaaggagtaaaatccatctcgccccaacttccc	1658
Db	4960	GAAAAGGGAGGGGTCCGAGGGTGTGTGGAGGATGAAATTCATCCGCCCCCACTTCCC	4901
Qy	1659	aaagagcagcatccctcccccgcgaaccaatgcaatgtctttaagttcaactctccgaagaa	1718
Db	4900	AAAGAGCAGCATCCCTCCCGGACCCATGATGTTTAAAGTCACTTTCCGAAGAGA	4841
Qy	1719	gtgaagaagttcaaggaacaatgtgcttgcaggtcccgagggagagcaatcaaaactcaca	1778
Db	4840	GTGAAGAGTTTCAAGACACTGTGCTTTCACAGCCCGGAGGACAGCATATCCAACTCACA	4781
Qy	1779	gaccagcaccctcttggagaacacggcctcttgcgcacaaactcaaggaacatcttgcg	1838
Db	4780	GACCAGCATCTCCCTTTTGAGACACCGCCTTCTTGCCACCATCTCACGGACACATTTTCTCC	4721
Qy	1839	ctagaanaacgctcttactagtcttctacatgtatgtgacatattcttaactaaagaatat	1898
Db	4720	CTAGAAACACACTTCTTACGTGCTCTTACATGATGATGGCATATCTTAACATAAACAATAT	4661
Qy	1899	tatttggggaaaaactcaagttgctgtacatatgctggaanaacttgcagaagcaataagct	1958
Db	4660	TATTGGGGGAAAAACTCAAGTGTGTGATATGCTGAGAACTGCAGAGCACTAATTAAGCT	4601
Qy	1959	gccaccccaanaactttttgaaatcatcttccagaanaaccttcaacttctgtagctt	2018

Db 4600 GCCACCCAAAATCTTTTGAATCATCTTCCAGACACCTCTTACTTCTGTAGTT 4541

QY 2019 ttaattgttaaaaaaaagttttaacagaagcacatgacatgaaagcctgcaga 2078

Db 4540 TTAAATGTTAAAAAAGTTTAAACAGACACATGACATATGAAGCCTGCAGGA 4481

QY 2079 ctgtgtcttttttgcacatcttcacgtgagcttgcacagaagaatgaatgagtg 2138

Db 4480 CTGTGTGTTTTTGGCAATCTTCCAGCTGGAGCTTGCACAGAAATGAAGTAGTGG 4421

QY 2139 ttcttaagaagtaagttacatattatttccacttaagttatttataagaaagttt 2198

Db 4420 TTTTAAAGAGTTAAGTTACATATTATTCTTCCACTTAAGTTATTATGCAAAAGTTT 4361

QY 2199 tctttagagaagatgacatgtaatatgtcttatagaatgaagtcgtcttccaga 2258

Db 4360 TCTTTAGAGAAATGACAAATGTAATATCTTTATGCAATTAACAGTCTGTTCCAGAG 4301

QY 2259 ttcagagacatgttataaagaacatgacatgcagcaag 2301

Db 4300 TCCAGAGACATTTTATTAAGACATGATCATGACCAAG 4258

LOCUS AC055813 151780 bp DNA HTG 22-JUL-2000

DEFINITION Homo sapiens chromosome 17 clone RP11-209M4 map 17, WORKING DRAFT

ACCESSION AC055813

VERSION AC055813.3 GI:9369521

KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 151780)

AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.

TITLE Homo sapiens chromosome 17, clone RP11-209M4

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 151780)

AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F., Boguslavsky,L., Boukhalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Dearlano,K., Dewar,K., Diaz,J.S., Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Grant,G., Hagos,B., Heatford,A., Horton,L., Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Laroque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheters,R., Melgrim,J., Meneus,L., Mihov,A., Miranda,C., Mlenga,V., Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N., Pissani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamaa,J., Testave,S., Theodore,J., Tirrell,A., Travers,M., Trigillo,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zalnoun,J., Zimmer,A. and Zody,M.

COMMENT

TITLE

JOURNAL

COMMENT

Submitted (18-APR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jul 22, 2000 this sequence version replaced g1:8920653.

ALL repeats were identified using RepeatMasker:

http://ftp.genome.washington.edu/RW/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence\_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L7949

Center clone name: 209\_M\_4

----- Summary Statistics

Sequencing vector: M13; M77815; 94% of reads

Sequencing vector: Plasmid; n/a; %0.1% of reads

6.158833063209080Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 139835 bases at least Q40

Consensus quality: 145751 bases at least Q30

Consensus quality: 148014 bases at least Q20

Insert size: 155000; agarose-fp

Insert size: 149480; sum-of-contigs

Quality coverage: 3.9 in Q20 bases; agarose-fp

Quality cov.

NOTE: This is a 'working draft' sequence. It currently consists of 24 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 1283: contig of 1283 bp in length

1284 1383: gap of 100 bp

1384 2931: contig of 1548 bp in length

2932 3031: gap of 100 bp

3032 4160: contig of 1129 bp in length

4161 4260: gap of 100 bp

4261 6377: contig of 2117 bp in length

6378 6477: gap of 100 bp

6478 8584: contig of 2107 bp in length

8585 8684: gap of 100 bp

8685 10824: contig of 2140 bp in length

10825 10924: gap of 100 bp

10925 13710: contig of 2786 bp in length

13711 13810: gap of 100 bp

13811 17465: contig of 3656 bp in length

17466 17566: gap of 100 bp

17567 19747: contig of 2181 bp in length

19748 19847: gap of 100 bp

19848 22997: contig of 3150 bp in length

22998 23097: gap of 100 bp

23098 26840: contig of 3743 bp in length

26841 26940: gap of 100 bp

26941 32331: contig of 5391 bp in length

32332 32431: gap of 100 bp

32432 37589: contig of 5158 bp in length

37590 37689: gap of 100 bp

37690 43814: contig of 6125 bp in length

43815 43914: gap of 100 bp

43915 49595: contig of 5681 bp in length

49596 49695: gap of 100 bp

49696 56982: contig of 7287 bp in length

56983 57082: gap of 100 bp

57083 63903: contig of 6821 bp in length

63904 64003: gap of 100 bp

64004 73106: contig of 9103 bp in length

73107 73206: gap of 100 bp

73207 81472: contig of 8266 bp in length

81473 81572: gap of 100 bp

81573 94653: contig of 13081 bp in length

94654 94753: gap of 100 bp

94754 106759: contig of 12006 bp in length

106760 106859: gap of 100 bp

106860 118633: contig of 11774 bp in length

118634 118733: gap of 100 bp

118734 135004: contig of 16271 bp in length

135005 135104: gap of 100 bp

135105 151780: contig of 16676 bp in length.

Location/Qualifiers

1..151780

FEATURES

Source



Db 113251 TCTGACAGCCAAAGATGAAACAAACAGAAAAAAGTAAGAGCTATTATGGCT 113310  
 QY 1479 gacatatttcggctgacaaactcccggaagaagctatgctctcccgagctgctcc 1538  
 Db 113311 GACATATTTCAGGCTGCAAACTCCGGAAGAGATGCTGCTCCAGCCGCTTCC 113370  
 QY 1539 ccgagatctgggttaactccaccctccatctcaagaataacatccatccatctgggta 1598  
 Db 113371 CCGGATTTGGGCTACTCCACCCTCCATCTCAAGAAATTAATCATCATTCATGGGGTA 113430  
 QY 1599 gaaaaggaagaaggtccgaggtgglyggagagagatagaatcacatccgcccaactccc 1658  
 Db 113431 GAAAAGGAGAGGGTCCGAGGGTGGAGAGGATAGAAATCATCCGCCCAACTTCCC 113490  
 QY 1659 aagaagagatccctcccccagccatagccagtgtttaagatccatccctccgaagaaga 1718  
 Db 113491 AAGAGAGAGATCCCTCCCGACCCATACGCTTTAAAGTCACTTCCTCCAAAGAGA 113550  
 QY 1719 gtgaaggttcaagagacatgagcttgcagcccgagagagagacatcacaaactcaca 1778  
 Db 113551 GTGAAGGTTCAAGAGACATGGCTTGACAGCCCGAGGGAGAGCATCAAAACTCACA 113610  
 1779 gaccagacatccctcttgagacacgcgctctgcccacacacacgcgagacatctctgc 1838  
 Db 113611 GACCAGACATCCCTTTTGAGACACGCGCTTCTGCCACACACACGACACATTTCTGC 113670  
 QY 1839 ctgagaagaagctctctactgctctacatgtagatgcatcttccactaaagaatat 1898  
 Db 113671 CTAGAAAACAGCTTCTTACTGCTTTCATGTATGATGCGATATCTTACTAAAGAAATAT 113730  
 QY 1899 tatctgggggaaaaaactaacaagtgctgtacatctgtagaagaactgcagagacataatgc 1958  
 Db 113731 TATGGGGGAAAAACTGACAGTGTCTGATGATGAGAACTGCAGAGCATATAATA-CT 113789  
 QY 1959 gccaccccaaaaactcttgaataatcatctccagacaaactctacttctgtgtgtt 2018  
 Db 113790 GCCACCCAAAAATCTTTTAAATCATTTCCAGACAACTCTTACTTCTGTGTGTGT 113849  
 QY 2019 ttaactgttaaaaaaaggtttaaacaagaagcacatgacatgaaagcctgcagaga 2078  
 Db 113850 TTATATTGTTAAAAAAGTTTAAACGAAGACATGACATGAAAGCCTGCAGGA 113909  
 QY 2079 ctggtcgcttcttggaactcttccacgtggagactgtccacaagaatgaagtagtg 2138  
 Db 113910 CTGCTCTTTTGGCAATCTTCCACGTGGGACTTGTCCAAAGATGAAGTGTGTG 113969  
 QY 2139 ttcttaagaaggttaagttacatattatcttccacttaagttatattgcaaaagttt 2198  
 Db 113970 TTTTAAAGGTTTAAATTTTATTTTCTCACTTAAGTTATTAAGCAAAAGTTT 114029  
 2199 tcttgtagaagaatgacatgttaataatgcttcttaataaataacagctcttctccagag 2258  
 Db 114030 TCTGTGAGAGATGACAAATGTTAATATGCTTATGATTAACAGTCTGTTCACAG 114089  
 QY 2259 tccaagagacatgttaataagaacatgaatcatgagccgaag 2301  
 Db 114090 TCCAGAGACATTTGTTAATAAGCAATGATCATGACCGAAG 114132

RESULT 7  
 AC073954/c DNA HTG 17-NOV-2000  
 LOCUS Homo sapiens chromosome 17 clone RP11-147C10, WORKING DRAFT  
 DEFINITION AC073954  
 ACCESSION AC073954  
 VERSION AC073954.3 GI:10998993  
 KEYWORDS HTG, HTGS\_PHASE1, HTGS\_DRAFT.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 177744)  
 AUTHORS Waterston, R.H.

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL

## COMMENT

The sequence of Homo sapiens clone  
 Unpublished  
 2 (bases 1 to 177744)  
 Waterston, R.H.  
 Direct Submission  
 Submitted (07-JUL-2000) Genome Sequencing Center, Washington  
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
 MO 63108, USA  
 On Oct 25, 2000 this sequence version replaced gi:9186649.

----- Genome Center -----  
 Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: <http://genome.wustl.edu/gsc/index.shtml>  
 ----- Project Information -----  
 Center project name: H\_NH0147C10  
 ----- Summary Statistics -----  
 Sequencing vector: M13, 100%  
 Sequencing vector: plasmid, 0%  
 Chemistry: Dye-Primer ET, 100% of reads  
 Chemistry: Dye-terminator Big Dye, 0% of reads  
 Assembly program: Phrap, version 0.990319  
 Consensus quality: 153679 bases at least Q40  
 Consensus quality: 163329 bases at least Q30  
 Consensus quality: 166780 bases at least Q20  
 Insert size: 185000; agarose-fp  
 Insert size: 183559; sum-of-contigs  
 Quality coverage: 3.29 in Q20 bases; agarose-fp  
 Quality coverage: 3.43 in Q20 bases; sum-of-contigs  
 -----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 34 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \*  
 \* 1 1247: contig of 1247 bp in length  
 \* 1248 1347: gap of unknown length  
 \* 1348 2762: contig of 1415 bp in length  
 \* 2763 2862: gap of unknown length  
 \* 2863 4773: contig of 1911 bp in length  
 \* 4774 4873: gap of unknown length  
 \* 4874 6326: contig of 1453 bp in length  
 \* 6327 6426: gap of unknown length  
 \* 6427 8873: contig of 2447 bp in length  
 \* 8874 8973: gap of unknown length  
 \* 8974 10783: contig of 1810 bp in length  
 \* 10784 10883: gap of unknown length  
 \* 10884 12763: contig of 1880 bp in length  
 \* 12764 12863: gap of unknown length  
 \* 12864 15381: contig of 2518 bp in length  
 \* 15382 15481: gap of unknown length  
 \* 15482 17365: contig of 1884 bp in length  
 \* 17366 17465: gap of unknown length  
 \* 17466 19992: contig of 2527 bp in length  
 \* 19993 20092: gap of unknown length  
 \* 20093 23512: contig of 3420 bp in length  
 \* 23513 23612: gap of unknown length  
 \* 23613 25699: contig of 2087 bp in length  
 \* 25700 25799: gap of unknown length  
 \* 25800 28333: contig of 2534 bp in length  
 \* 28334 28433: gap of unknown length  
 \* 28434 32027: contig of 3594 bp in length  
 \* 32028 32127: gap of unknown length  
 \* 32128 36115: contig of 3988 bp in length  
 \* 36116 36215: gap of unknown length  
 \* 36216 39882: contig of 3667 bp in length  
 \* 39883 39983: gap of unknown length  
 \* 39984 44009: contig of 4027 bp in length  
 \* 44109 44110: gap of unknown length  
 \* 47829: contig of 3720 bp in length



*	4780	47929:	gap of unknown length
*	47930	53768:	contig of 5839 bp in length
*	53769	53868:	gap of unknown length
*	53869	60215:	contig of 6347 bp in length
*	60216	60315:	gap of unknown length
*	60316	66477:	contig of 6162 bp in length
*	66478	66577:	gap of unknown length
*	66578	71476:	contig of 4899 bp in length
*	71477	71575:	gap of unknown length
*	71577	78883:	contig of 7307 bp in length
*	78884	78963:	gap of unknown length
*	78984	85952:	contig of 6969 bp in length
*	85953	86052:	gap of unknown length
*	86053	93510:	contig of 7458 bp in length
*	93511	93610:	gap of unknown length
*	93610	100718:	contig of 7108 bp in length
*	100719	100818:	gap of unknown length
*	100819	109785:	contig of 8977 bp in length
*	109796	109855:	gap of unknown length
*	109866	117780:	contig of 7895 bp in length
*	117791	117890:	gap of unknown length
*	117891	126044:	contig of 8154 bp in length
*	126045	126144:	gap of unknown length
*	126145	134442:	contig of 8298 bp in length
*	134443	134542:	gap of unknown length
*	134543	145763:	contig of 11121 bp in length
*	145764	157189:	gap of unknown length
*	157190	157289:	contig of 11426 bp in length
*	157290	166822:	gap of unknown length
*	166823	166932:	contig of 9533 bp in length
*	166923	177744:	contig of 10822 bp in length

## FEATURES

**source**

misc_feature	1. 1247
misc_feature	/note=assembly_name:Contig26
misc_feature	1346. 1262
misc_feature	/note=assembly_name:Contig31
misc_feature	2863. 4773
misc_feature	/note=assembly_name:Contig34
misc_feature	4874. 6326
misc_feature	/note=assembly_name:Contig35
misc_feature	6427. 8873
misc_feature	/note=assembly_name:Contig36
misc_feature	8974. 10783
misc_feature	/note=assembly_name:Contig37
misc_feature	10884. 12763
misc_feature	/note=assembly_name:Contig38
misc_feature	12864. 15381
misc_feature	/note=assembly_name:Contig39
misc_feature	15462. 17365
misc_feature	/note=assembly_name:Contig40
misc_feature	17466. 19992
misc_feature	/note=assembly_name:Contig41
misc_feature	20093. 23512
misc_feature	/note=assembly_name:Contig42
misc_feature	23613. 25699
misc_feature	/note=assembly_name:Contig43
misc_feature	25800. 28333
misc_feature	/note=assembly_name:Contig44
misc_feature	28434. 32027
misc_feature	/note=assembly_name:Contig45
misc_feature	32128. 36115
misc_feature	/note=assembly_name:Contig46
misc_feature	36216. 39882
misc_feature	/note=assembly_name:Contig47
misc_feature	39983. 44009
misc_feature	/note=assembly_name:Contig48
misc_feature	44110. 47829

```
misc_feature /note="assembly_name:Contig49" 47930..53768 /note="assembly_name:Contig50" 53869..60215 /note="assembly_name:Contig51" 60316..66477 /note="assembly_name:Contig52" 66578..71476 /note="assembly_name:Contig53" 71577..78883 /note="assembly_name:Contig54" 78984..85952 /note="assembly_name:Contig55 clone_end:47 vector_side:left" 86053..93510 /note="assembly_name:Contig56" 93611..100718 /note="assembly_name:Contig57" 100819..109795 /note="assembly_name:Contig58" 109896..117790 /note="assembly_name:Contig59" 117891..126044 /note="assembly_name:Contig60" 126145..134442 /note="assembly_name:Contig61" 134543..145663 /note="assembly_name:Contig62" 145764..157189 /note="assembly_name:Contig63" 157290..166822 /note="assembly_name:Contig64" 166923..177744 /note="assembly_name:Contig65" 177845..188965 /note="assembly_name:Contig66" 189066..198186 /note="assembly_name:Contig67" 198287..207407 /note="assembly_name:Contig68" 207508..216628 /note="assembly_name:Contig69" 216729..225849 /note="assembly_name:Contig70" 225950..235070 /note="assembly_name:Contig71" 235171..244291 /note="assembly_name:Contig72" 244392..253512 /note="assembly_name:Contig73" 253613..262733 /note="assembly_name:Contig74" 262834..271954 /note="assembly_name:Contig75" 272055..281175 /note="assembly_name:Contig76" 281276..290396 /note="assembly_name:Contig77" 290497..300000 /note="assembly_name:Contig78" 300101..309221 /note="assembly_name:Contig79" 309322..318442 /note="assembly_name:Contig80" 318543..327663 /note="assembly_name:Contig81" 327764..336884 /note="assembly_name:Contig82" 336985..346105 /note="assembly_name:Contig83" 346206..355326 /note="assembly_name:Contig84" 355427..364547 /note="assembly_name:Contig85" 364648..373768 /note="assembly_name:Contig86" 373869..382989 /note="assembly_name:Contig87" 383090..392210 /note="assembly_name:Contig88" 392311..401431 /note="assembly_name:Contig89" 401532..410652 /note="assembly_name:Contig90" 410753..419873 /note="assembly_name:Contig91" 419974..429094 /note="assembly_name:Contig92" 429195..438315 /note="assembly_name:Contig93" 438416..447536 /note="assembly_name:Contig94" 447637..456757 /note="assembly_name:Contig95" 456858..465978 /note="assembly_name:Contig96" 466079..475199 /note="assembly_name:Contig97" 475300..484420 /note="assembly_name:Contig98" 484521..493641 /note="assembly_name:Contig99" 493742..502862 /note="assembly_name:Contig100" 502963..512083 /note="assembly_name:Contig101" 512184..521304 /note="assembly_name:Contig102" 521405..530525 /note="assembly_name:Contig103" 530626..539746 /note="assembly_name:Contig104" 539847..548967 /note="assembly_name:Contig105" 549068..558188 /note="assembly_name:Contig106" 558289..567409 /note="assembly_name:Contig107" 567510..576630 /note="assembly_name:Contig108" 576731..585851 /note="assembly_name:Contig109" 585952..595072 /note="assembly_name:Contig110" 595173..604293 /note="assembly_name:Contig111" 604394..613514 /note="assembly_name:Contig112" 613615..622735 /note="assembly_name:Contig113" 622836..631956 /note="assembly_name:Contig114" 632057..641177 /note="assembly_name:Contig115" 641278..650398 /note="assembly_name:Contig116" 650499..659619 /note="assembly_name:Contig117" 659720..668840 /note="assembly_name:Contig118" 668941..678061 /note="assembly_name:Contig119" 678162..687282 /note="assembly_name:Contig120" 687383..696503 /note="assembly_name:Contig121" 696604..705724 /note="assembly_name:Contig122" 705825..714945 /note="assembly_name:Contig123" 715046..724166 /note="assembly_name:Contig124" 724267..733387 /note="assembly_name:Contig125" 733488..742608 /note="assembly_name:Contig126" 742709..751829 /note="assembly_name:Contig127" 751930..761050 /note="assembly_name:Contig128" 761151..770271 /note="assembly_name:Contig129" 770372..779492 /note="assembly_name:Contig130" 779593..788713 /note="assembly_name:Contig131" 788814..797934 /note="assembly_name:Contig132" 798035..807155 /note="assembly_name:Contig133" 807256..816376 /note="assembly_name:Contig134" 816477..825597 /note="assembly_name:Contig135" 825698..834818 /note="assembly_name:Contig136" 834919..844039 /note="assembly_name:Contig137" 844140..853260 /note="assembly_name:Contig138" 853361..862481 /note="assembly_name:Contig139" 862582..871702 /note="assembly_name:Contig140" 871803..880923 /note="assembly_name:Contig141" 881024..890144 /note="assembly_name:Contig142" 890245..899365 /note="assembly_name:Contig143" 899466..908586 /note="assembly_name:Contig144" 908687..917807 /note="assembly_name:Contig145" 917908..927028 /note="assembly_name:Contig146" 927129..936249 /note="assembly_name:Contig147" 936350..945470 /note="assembly_name:Contig148" 945571..954691 /note="assembly_name:Contig149" 954792..963912 /note="assembly_name:Contig150" 964013..973133 /note="assembly_name:Contig151" 973234..982354 /note="assembly_name:Contig152" 982455..991575 /note="assembly_name:Contig153" 991676..1000796 /note="assembly_name:Contig154" 1000897..1009917 /note="assembly_name:Contig155" 1010018..1019138 /note="assembly_name:Contig156" 1019239..1028359 /note="assembly_name:Contig157" 1028460..1037580 /note="assembly_name:Contig158" 1037681..1046801 /note="assembly_name:Contig159" 1046902..1056022 /note="assembly_name:Contig160" 1056123..1065243 /note="assembly_name:Contig161" 1065344..1074464 /note="assembly_name:Contig162" 1074565..1083685 /note="assembly_name:Contig163" 1083786..1092906 /note="assembly_name:Contig164" 1093007..1102127 /note="assembly_name:Contig165" 1102228..1111348 /note="assembly_name:Contig166" 1111449..1120569 /note="assembly_name:Contig167" 1120670..1129790 /note="assembly_name:Contig168" 1129891..1139011 /note="assembly_name:Contig169" 1139112..1148232 /note="assembly_name:Contig170" 1148333..1157453 /note="assembly_name:Contig171" 1157554..1166674 /note="assembly_name:Contig172" 1166775..1175895 /note="assembly_name:Contig173" 1175996..1185116 /note="assembly_name:Contig174" 1185217..1194337 /note="assembly_name:Contig175" 1194438..1203558 /note="assembly_name:Contig176" 1203659..1212779 /note="assembly_name:Contig177" 1212880..1222000 /note="assembly_name:Contig178" 1222101..1231221 /note="assembly_name:Contig179" 1231322..1240442 /note="assembly_name:Contig180" 1240543..1249663 /note="assembly_name:Contig181" 1249764..1258884 /note="assembly_name:Contig182" 1258985..1268105 /note="assembly_name:Contig183" 1268206..1277326 /note="assembly_name:Contig184" 1277427..1286547 /note="assembly_name:Contig185" 1286648..1295768 /note="assembly_name:Contig186" 1295869..1304989 /note="assembly_name:Contig187" 1305090..1314210 /note="assembly_name:Contig188" 1314311..1323431 /note="assembly_name:Contig189" 1323532..1332652 /note="assembly_name:Contig190" 1332753..1341873 /note="assembly_name:Contig191" 1341974..1351094 /note="assembly_name:Contig192" 1351195..1360315 /note="assembly_name:Contig193" 1360416..1369536 /note="assembly_name:Contig194" 1369637..1378757 /note="assembly_name:Contig195" 1378858..1387978 /note="assembly_name:Contig196" 1388079..1397199 /note="assembly_name:Contig197" 1397300..1406420 /note="assembly_name:Contig198" 1406521..1415641 /note="assembly_name:Contig199" 1415742..1424862 /note="assembly_name:Contig200" 1424963..1434083 /note="assembly_name:Contig201" 1434184..1443304 /note="assembly_name:Contig202" 1443405..1452525 /note="assembly_name:Contig203" 1452626..1461746 /note="assembly_name:Contig204" 146184
```

Query Match	46.3%	Score 1066.4	DB 75	Length 177744
Best Local Similarity	97.0%	Pred. No. 5.6e-195		
Matches 1129	Conservative	0	Mismatches 31	Indels 4
				Gaps 4
QY 1142	aactgtagatggtgttct-agtctgtgctctgcac- taactgtcgtgtaaccttgaa	1199		
Db 93369	AGCTGGAGTGTGTGTGTCAGTACCTCGGCTGCACCTTACTCTGTGGGAATCCGGAA	93310		
QY 1200	ctaacattctctcttcggacctaaattcca-cttggtaaatgaggttgaggtgg	1258		
Db 93309	CTAACAGTCTGCTGGCGGGGACTCGATTACTCATGTGTAAATGTATGGTGAAGGCGG	93250		
QY 1259	aatgagcttcggagagactattgataatgattlccaaagactcaatgtgctttgaatg	1318		
Db 93249	AATGAGTCTCGAGGAGACTATTGGCATGTAGTCCAAAGSACTCCAGTGGCTTTTGAATG	93190		
QY 1319	ggcagaggttaagagaagaagagaagaagaagaatgaatgaggtgcttgaattcat	1378		
Db 93189	GGCGAGAGTATGAGAGAGAGAGAGAGAGAGAGATGAATATACAGTTCAATTATTCAG	93130		
QY 1379	gccaaagtctactccagaatt- cagaattgtatgtctctcttcgaaagccaaagaatgaa	1437		
Db 93129	GCCAGGTCCTCTCCAGAAATTNACAGTGTGATCTCTCTTTCGACAGCCCAAGATGAA	93070		
QY 1438	aaacaaacagaaaaaaagaagtaagagatcattatagctgagcaatttaagcgtgaca	1497		
Db 93069	AAACAAACAGAAAAAAGTAAGAGCTATTATTATGCGTGAATATTTCAGCGCTGACA	93010		
QY 1498	aaactctgagaagaagctatgtctgtccagagctgtgcttcccgagatglttgctaccto	1557		
Db 93009	AATCTCTGGAAGAGCATGTCTGCTTCCAGCGCTGGCTTCCCGAGTTTGGCTCTACAC	92950		
QY 1558	cacctccatctccaagaataaacatctccattgaggtgagaaagagaggttcgag	1617		

Db 92949 CACCCCTCATCTCAAGAATAACATCATCCATTGGGGTAGAAAAAGAGAGGCTCCGAG 92890  
 Qy 1618 ggtggtgagggagatagaataacatccgcccccaactccccaaagagagatccccc 1677  
 Db 92889 GGTGGTGGGGGATAGAAATCATCCGCCCACTTCCCAAAAGAGAGATCCCTCC 92830  
 Qy 1678 ccgaccatagcatgttttaaaagtcaccttcgagaagaagtgaaaagttcaagagcac 1737  
 Db 92829 CCGACCATAGCATGTTTAAAGTCACCTTCGGAAGAGAGTGAAGGTTCAAGGACAC 92770  
 Qy 1738 tggccttgagggagggagagagacgacatcaaaaactacagagacacatccctttg 1797  
 Db 92769 TGGCTTGGAGGCCGAGGAGAGGACATCAAACTCAGACACGACATCCCTTTG 92710  
 Qy 1798 agacaccgcttcgcccacacacacacagagacatcttcgctaagaacagctcttac 1857  
 Db 92709 AGACACGCGCTTCGCCACACACTCAGGACATTTCTGCCTAGAAAACAGCTTCCTAC 92650  
 Qy 1858 tgccttaacatgtagatgagatatacttaacataaagaatattatgggggaaaaactaca 1917  
 Db 92649 TGCTTACATGTGATGAGCATATCTTACACTAAAGAAATATTATTTGGGGAAAAACTACA 92590  
 Qy 1918 agtgcgtacatatgctggaagacatgacagataatagctgcacacccaacatctttt 1977  
 Db 92589 AGTGCCTACATATGCTGAGAAACCTCAGAGCATATCTGCGACCTCAAAAATCTTTT 92530  
 Qy 1978 gaaatcatttcagagaacaccttaactctctgtgtagttttaaattgtaaaaaaaa 2037  
 Db 92529 GAAATCATTTCCAGAACCTCTTACCTTGTGTACTTTTAAATGTTTAAAAAAA 92470  
 Qy 2038 agtttaaacagaagacacatgacatatagaagacgagagatgctgttcttttgcaa 2097  
 Db 92469 AGTTTAAACAGAACATGACATATGAAACCTCAGAGCATGCTGCTTTTGTGCAA 92410  
 Qy 2098 ttcttcacgtgagactgttccaaagaatgaagtaagtgtttttaaagaagtaagta 2157  
 Db 92409 TTTCTCCACGTGGACTTGTCCACAAGAAATGAAAGTACTGTTTAAAGGTTAAAGTTA 92350  
 Qy 2158 catactatcttcacactaaagtattatgcaaaagcttctctgtgaagaatgaacat 2217  
 Db 92349 CATATTTATTTTCTCACTTAAGTTATTATGCAAAAGTTTCTGTGATGAAATGACAT 92290  
 Qy 2218 gttcaatctgcttaagaatcaagctcgtcttcacagagtcagagacatgttata 2277  
 Db 92289 GTTAATATTCCTTATGAATTAACAAGTCTTCTTCAGAGTCAGAGACATTTGTATA 92230  
 Qy 2278 aagacaatgaatcatgaccgaag 2301  
 Db 92229 AAGCAATGAAATCATGACCGAAAG 92206  
 RESULT 8  
 AF326742 642 bp mRNA PRI 28-FEB-2001  
 DEFINITION Cercopithecus aethiops sclerostlin mRNA, complete cds.  
 ACCESSION AF326742  
 VERSION AF326742.1 GI:13161028  
 KEYWORDS  
 SOURCE African green monkey.  
 ORGANISM Cercopithecus aethiops.  
 Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;  
 Cercopithecinae; Cercopithecus.  
 REFERENCE 1 (bases 1 to 642)  
 AUTHORS Brunkow, M.E., Gardner, J.C., Van Ness, J., Paepker, B.W.,  
 Kovacevich, B.R., Prohl, S., Skonier, J.E., Zhao, L., Sabo, P.J.,  
 Fu, Y.H., Altsch, R.S., Gillett, L., Colbert, T., Tacconi, P., Galas, D.,  
 Hamersma, H., Beighton, P., and Mulligan, J.T.  
 TITLE Bone Dysplasia Sclerosteosis Results from Loss of the SOST Gene  
 Product, a Novel Cysteine Knot-Containing Protein  
 JOURNAL Am. J. Hum. Genet. 68 (3), 577-589 (2001)  
 PUBMED 11179006  
 REFERENCE 2 (bases 1 to 642)

AUTHORS Brunkow, M.E., Gardner, J.C., Van Ness, J., Paepker, B.W.,  
 Kovacevich, B.R., Prohl, S., Skonier, J.E., Zhao, L., Sabo, P.J.,  
 Fu, Y.H., Altsch, R.S., Gillett, L., Colbert, T., Tacconi, P., Galas, D.,  
 Hamersma, H., Beighton, P., and Mulligan, J.T.  
 TITLE Direct Submission  
 JOURNAL Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 1631  
 220th St. SE, Bothell, WA 98021, USA  
 FEATURES  
 source location/Qualifiers  
 1. .642  
 /organism="Cercopithecus aethiops"  
 /db\_xref="taxon:9534"  
 1. .642  
 /product="sclerostlin"  
 /protein\_id="AAK13457.1"  
 /db\_xref="GI:13161029"  
 /translation="MQLPLALCLVCLLVAHAFRYVEGGMOAFKNDATETILPELGEYP  
 EPPPELENKNTNRAENGRPHHPFERKDVSEVSGRELTHTRYTVDGPPRSAPKPYE  
 LVSGCGGPARLLPNAIGRKMMRRSGDFPCIDPRYRAQVOLCLCCGAAAPRAKRYR  
 LVASCKCKRLTRFHNSLEKDRGPEEARFQGRKFRPRARAKANQALENAI"  
 BASE COUNT 113 a 228 c 210 g 91 t  
 ORIGIN  
 Query Match 26.5% Score 610: DB 89: Length 642:  
 Best Local Similarity 96.9%: Pred. No. 4.1e-107:  
 Matches 622: Conservative 0: Mismatches 20: Indels 0: Gaps 0:  
 Qy 48 atcgagctccacatgagcctgtgtctatctgtctgtgtacacacagccttcgltga 107  
 Db 1 ATGCACTCCCACTAGCGCCGCTGTCTGTCTGCTGCTGTGTACACGACCTTCCGTGTA 60  
 Qy 108 gtggagggccaggggtgtggagggcggtcaagaatgagtcagagaaatcccgagagtc 167  
 Db 61 GTGGAGGGCCAGGGGTGTGGAGGCTTCAAGATGATGACCAAGAAATATCTCCCAAGCTC 120  
 Qy 168 ggaagagtaaccccgagcctccacccgagagcttggaagaaacaaagaccatgaacccggcgag 227  
 Db 121 GGAGAGTACCCCGAGGCTCCACCGAGACTGGAAGAACAAAGACATGAAACCGGCGGAG 180  
 Qy 228 aacggaggggggctcccaacaccccttgagacaaagagctgtccgagtagagctgc 287  
 Db 181 AATGAGGGGGGCGCTCCCGACACCCCTTGTGAGAACCAAGACGTGTCCAGTACAGCTGC 240  
 Qy 288 cgcgagctgcaatccacccgactacagtgacagatgggacgtgacgagcgccaaagcggtc 347  
 Db 241 CGAGAGCTGCACCTTACCCCGCTACGATGACGATGGGCGCTGCGGACGCCCAAGCGATGC 300  
 Qy 348 accgagctgtgtgtcttcgagcagatgagcgagcgagcgctctgtgccaagacacatgcgc 407  
 Db 301 ACCGAGTTGTGTCTCCGGCCAGTGGCGGCGCCGACGCTGTCTGCCAAGCCATTCGGC 360  
 Qy 408 cgcggaagtggtgagacactagtgagccgagcttcacacacacagtcgagacacgctaccgc 467  
 Db 361 CGCGCAAGTGTGTGGCGCCGAGTGGGCGCGGACTTCCGTCATCCCGACGCTAACGC 420  
 Qy 468 ggcgagcgctgagcagctgt 527  
 Db 421 GCGCAGCTGTGACAGCTGT 480  
 Qy 528 ctggtgagctgtgcaagtgcaagcgctcaccgcttcacacacacagtcgagagtcag 587  
 Db 481 CTGGTGGCTGTGTCAAGTGAAGCGGCTTCAACCGCTTCCACACAGTGGAGCTCAAG 540  
 Qy 588 gactcggagacgagcgctctggcgagaaagagcgccgagcccgccgagcgcg 647  
 Db 541 GACTTGGTCCCGAGGCGGCTGCGCCGCAAAAGCGGGAACCGCGGCGCCGCGCGG 600  
 Qy 648 agcgcaagacgaagcgagcgtgtgagaaagcctactag 689  
 Db 601 GGGGCAAGGCAATCAAGCGCGAGCTGAGAAAGCGCTACTAG 642

RESULT	9
AF326741	AF326741 674 bp mRNA ROD 28-FEB-2001
LOCUS	Rattus norvegicus sclerostin mRNA, complete cds.
DEFINITION	AF326741 Rattus norvegicus sclerostin mRNA, complete cds.
ACCESSION	AF326741.1 GI:13161025
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	Norway rat. Rattus norvegicus Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE	1 (bases 1 to 674) Brunkow M.E., Gardner J.C., Van Ness, J., Paepke B.W., Kovacevich B.R., Prolli S., Skoner J.E., Zhao L., Sabo P.J., Fu Y.H., Alish R.S., Gillett L., Colbert T., Tacconi P., Galas D., Hamersta H., Beighton P., Mulligan J.T. Bone Dysplasia Sclerosteosis Results from Loss of the SOST Gene Product, a Novel Cystine Knot-Containing Protein Am. J. Hum. Genet. 68 (3), 577-589 (2001)
JOURNAL	11179006 2 (bases 1 to 674) Brunkow M.E., Gardner J.C., Van Ness, J., Paepke B.W., Kovacevich B.R., Prolli S., Skoner J.E., Zhao L., Sabo P.J., Fu Y.H., Alish R.S., Gillett L., Colbert T., Tacconi P., Galas D., Hamersta H., Beighton P., Mulligan J.T. Direct Submission Submitted (07-Dec-2000) Genomics, Celltech Chromosome Inc., 1631 220th St. SE, Bothell, WA 98021, USA
TITLE	Location/Qualifiers
AUTHORS	1..674 /organism="Rattus norvegicus" /strain="Sprague-Dawley" /db_xref="taxon:10116"
FEATURES	5'UTR 1..32 CDS 33..674 /codon_start=1 /product="sclerostin" /protein_id="AAK13456.1" /db_xref="GI:13161026" translation="MQSLAPCLACL VHAFAVAVESQGAERKNDATETII PGLEVRV EPDLENNQTMRNENGRRPHPIPTDKSVSYSRRLRYFVDGPRCAKVTVE IVASCCGPARIILPNAIGRVKMWRPNPDRCIPDDEYRAQRVOLLCPGAARSKVA LVAASKCRILRFNPNSSELDKFGEZTRYPQKGKRPPRARAKAMDAELENAY"
BASE COUNT	128 a 240 c 208 g 98 t
ORIGIN	
Query Match:	23.2% Score 533; DB 94; Length 674;
Best Local Similarity	87.9%; Pred. No. 2,46-92;
Matches 581; Conservative	0; Mismatches 80; Indels 0; Gaps 0;
Oy	29 cccctcctgctgctgtacatgaactcccaactggcccccttgctcatcctgcctgctgct 88
Dd	14 CCTTCCTCTTTCGGAGCACCATGCACCTGCATACCCCTTGCTTCCTCGCTGTG 73
Oy	89 acacaagaccttcgcygtagttagaaggccaaggggtggaagcgcttcaaagaatgatccac 148
Dd	74 ACATGCAAGCCTTGTTGCTGCTGTGGAAGGCCAAGGGGTGGAGAACCCTTCAAGAAATATGCCCAC 133
Oy	149 ggaaatatcctcgagagcttcgsagatgcccgagcttcacacggagcttgagagaaca 208
Dd	134 AGAATCATATCCGGGACTCAGAGAGTACCCACAAGCCTTCCTCAGGAACCTTAGAACAACA 193
Oy	209 gaaccatgaaccggcgagaaacgagggcgagctcccaccaaacccctttgagaccaaaga 268
Dd	194 GACCATGAACCGGGCGAGAACGAGAGAGAGACCCTCCACCATCTTATGACACCAAAGA 253
Oy	269 cgctcgcgaataagctgcgcggaactgaacttaaccccgcaatggacggatggcgatggcg 328
Dd	254 CGTGTCCGATTAAGCTGGCCGGAGACTGCATCAACACCGGCTTCCTGATGACAGACGCCCG 313
Oy	329 ccgcagcgcaagcggtgctaccagagctgtgtgtgtctcgcgagctgcgagcccgagcgccct 388

	Db	314	CCGAGTGGCCAAACCGGTACCAGATTGTTGCTCGAGGCCAATGGCGCCCGCGCGGT	373
	Qy	389	gctgcaccaacgcatcgtcgccgcgaagtgttgagaaactatgttggcccaattccgtg	448
	Db	374	GCTGCCCAACCGCATGGAGCGGTAAGTGTTGGCGCCCGCAACGACCCGACTTCGC	433
	Qy	449	cataccgaacgcgtaccgcgcgcgaacgagtcgttcagaatcgtgtttcgcgttgttagagcgc	508
	Db	434	CATCCGGAATGCTACCGCGCGCACCGGGTACAGTGTGTGGCCCCGCGCGCGCGC	493
	Qy	509	ggcgcgcgcgaagatgtgtcctgtgtccttcgttaaatgtgaacgcgtaccgcgttcaa	568
	Db	494	GGCTCGCGCAAGTGTGCTGTGTGTGGCTCTGTGCAATGCAAGGAGGCTCAACCGCTTCCA	553
	Qy	569	caaacagtcgagatcaaggacttcggagaccgaagcgcgtctgcgcgaagaagcgcgaa	628
	Db	554	CAACCAATCGAGCTTCMAAGACTTCGGAGCTGAGACCGCGCGCGCGAGAAAGGTCCGA	613
	Qy	629	ggcgcgcgcgcgcgcgcgcgcgcgcgaagccaacgaacgaacgcgaagctgtgaagaccta	688
	Db	614	GGCGGCGCCCGCGCGCGGGAGCCAAAGCCAACACGAGCGAGATGGAAGAAGCTTACTA	673
	Qy	689	g 689	
	Db	674	G 674	
	RESULT	10		
	AC073954			
	LOCUS			
	DEFINITION	AC073954 bp DNA HTG 17-NOV-2000		
	ACCESSION	Homo sapiens chromosome 17 clone RP11-147C10, WORKING DRAFT		
	VERSION	AC073954.3 GI:10989993		
	SOURCE	HTG, HTGS_PHASE1; HTGS_DRAFT.		
	ORGANISM	human.		
	REFERENCE	Homo sapiens		
	AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.		
	JOURNAL	1 (bases 1 to 17744)		
	REFERENCE	The sequence of Homo sapiens clone		
	AUTHORS	Unpublished		
	TITLE	2 (bases 1 to 17744)		
	JOURNAL	Waterston R.H		
	REFERENCE	Direct Submission		
	AUTHORS	University (07-JUL-2000) Genome Sequencing Center, Washington		
	TITLE	Submitted School of Medicine, 4444 Forest Park Parkway, St. Louis,		
	JOURNAL	MO 61(08), USA		
	COMMENT	On Oct 25, 2000 this sequence version replaced gi:9186849.		
		----- Genome Center -----		
		Center: Washington University Genome Sequencing Center		
		Center code: WUGSC		
		Web site:http://genome.wustl.edu/gsc/index.shtml		
		----- Project Information -----		
		Center project name: H_NH0147C10		
		----- Summary Statistics -----		
		Sequencing vector: M13; 100%		
		Sequencing vector: plasmid; 0%		
		Chemistry: Dye-primer ET; 100% of reads		
		Chemistry: Dye-terminator Big Dye; 0% of reads		
		Assembly program: Phrap; version 0.990319		
		Consensus quality: 155679 bases at least Q40		
		Consensus quality: 163329 bases at least Q30		
		Consensus quality: 166780 bases at least Q20		
		Insert size: 185000; agarose-fp		
		Insert size: 183559; sum-of-contigs		
		Quality coverage: 3.29 in Q20 bases; agarose-fp		
		Quality coverage: 3.43 in Q20 bases; sum-of-contigs		
		* NOTE: This is a working draft sequence. It currently		

\* consists of 34 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1247: contig of 1247 bp in length  
\* 1248 1347: gap of unknown length  
\* 1348 2762: contig of 1415 bp in length  
\* 2763 2862: gap of unknown length  
\* 2863 4773: contig of 1911 bp in length  
\* 4774 4873: gap of unknown length  
\* 4874 6326: contig of 1453 bp in length  
\* 6327 6426: gap of unknown length  
\* 6427 8873: contig of 2447 bp in length  
\* 8874 8973: gap of unknown length  
\* 8974 10783: contig of 1810 bp in length  
\* 10784 10883: gap of unknown length  
\* 10884 12763: contig of 1880 bp in length  
\* 12764 12863: gap of unknown length  
\* 12864 15381: contig of 2518 bp in length  
\* 15382 15481: gap of unknown length  
\* 15482 17365: contig of 1884 bp in length  
\* 17366 17465: gap of unknown length  
\* 17466 19992: contig of 2527 bp in length  
\* 19993 20092: gap of unknown length  
\* 20093 23512: contig of 3420 bp in length  
\* 23513 23612: gap of unknown length  
\* 23613 25699: contig of 2087 bp in length  
\* 25700 28333: contig of 2534 bp in length  
\* 28334 28433: gap of unknown length  
\* 28434 32027: contig of 3594 bp in length  
\* 32028 32127: gap of unknown length  
\* 32128 36115: contig of 3988 bp in length  
\* 36116 36215: gap of unknown length  
\* 36216 39882: contig of 3667 bp in length  
\* 39883 39983: gap of unknown length  
\* 39984 44009: contig of 4027 bp in length  
\* 44010 47829: gap of unknown length  
\* 47830 47929: gap of unknown length  
\* 47930 53768: contig of 5839 bp in length  
\* 53769 53868: gap of unknown length  
\* 53869 60215: contig of 6347 bp in length  
\* 60216 60315: gap of unknown length  
\* 60316 66477: contig of 6162 bp in length  
\* 66478 66577: gap of unknown length  
\* 66578 71476: contig of 4899 bp in length  
\* 71477 71576: gap of unknown length  
\* 71577 78883: contig of 7307 bp in length  
\* 78884 85952: gap of unknown length  
\* 85953 86052: contig of 6969 bp in length  
\* 86053 93510: contig of 7458 bp in length  
\* 93511 93610: gap of unknown length  
\* 93611 100718: contig of 7108 bp in length  
\* 100719 100818: gap of unknown length  
\* 100819 109795: contig of 8977 bp in length  
\* 109796 109895: gap of unknown length  
\* 109896 117790: contig of 7895 bp in length  
\* 117791 117890: gap of unknown length  
\* 117891 126044: contig of 8154 bp in length  
\* 126045 126144: gap of unknown length  
\* 126145 134442: contig of 8298 bp in length  
\* 134443 134542: gap of unknown length  
\* 134543 145663: contig of 11121 bp in length  
\* 145664 145763: gap of unknown length  
\* 145764 157189: contig of 11426 bp in length  
\* 157190 157289: gap of unknown length  
\* 157290 166822: contig of 9533 bp in length  
\* 166823 166922: gap of unknown length

FEATURES \* 166923 177744: contig of 10822 bp in length.  
source  
1. 177744  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="17"  
/clone="RP11-147C10"  
1. 1247  
misc\_feature  
/note="assembly\_name:Contig26"  
1348..2762  
misc\_feature  
/note="assembly\_name:Contig31"  
2863..4773  
misc\_feature  
/note="assembly\_name:Contig34"  
4874..6326  
misc\_feature  
/note="assembly\_name:Contig35"  
6427..8873  
misc\_feature  
/note="assembly\_name:Contig36"  
8974..10783  
misc\_feature  
/note="assembly\_name:Contig37"  
10884..12763  
misc\_feature  
/note="assembly\_name:Contig38"  
12864..15381  
misc\_feature  
/note="assembly\_name:Contig39"  
15482..17365  
misc\_feature  
/note="assembly\_name:Contig40"  
17466..19992  
misc\_feature  
/note="assembly\_name:Contig41"  
20093..23512  
misc\_feature  
/note="assembly\_name:Contig42"  
23613..25699  
misc\_feature  
/note="assembly\_name:Contig43"  
25800..28333  
misc\_feature  
/note="assembly\_name:Contig44"  
28434..32027  
misc\_feature  
/note="assembly\_name:Contig45"  
32128..36115  
misc\_feature  
/note="assembly\_name:Contig46"  
36216..39882  
misc\_feature  
/note="assembly\_name:Contig47"  
39983..44009  
misc\_feature  
/note="assembly\_name:Contig48"  
44110..47829  
misc\_feature  
/note="assembly\_name:Contig49"  
47930..53768  
misc\_feature  
/note="assembly\_name:Contig50"  
53869..60215  
misc\_feature  
/note="assembly\_name:Contig51"  
60316..66477  
misc\_feature  
/note="assembly\_name:Contig52"  
66578..71476  
misc\_feature  
/note="assembly\_name:Contig53"  
71577..78883  
misc\_feature  
/note="assembly\_name:Contig54"  
78984..85952  
misc\_feature  
/note="assembly\_name:Contig55"  
clone\_end:17  
vector\_side:left  
misc\_feature  
/note="assembly\_name:Contig56"  
93611..100718  
misc\_feature  
/note="assembly\_name:Contig57"  
100819..109795  
misc\_feature  
/note="assembly\_name:Contig58"  
109896..117790  
misc\_feature  
/note="assembly\_name:Contig59"  
117891..126044  
misc\_feature  
/note="assembly\_name:Contig60"  
126145..134442  
misc\_feature  
/note="assembly\_name:Contig61"  
134543..145663  
misc\_feature  
/note="assembly\_name:Contig62"  
145764..157189  
misc\_feature  
/note="assembly\_name:Contig63"



[illegible]

Accession	Gene	Protein	Species	Length (aa)	Weight (kDa)	PI	Source
316	ccgcttgagccgctgcgcgaagcgcgaagcgcggtaccggaacttgctgtctcggccagatg		Human	375	42.5	5.2	GenBank
178	CCGATGGGCGCGTCCCGAGGCCAAGCCGCGTACCAGACTGTGTGCTCGGCGCAATGCC		Human	237	27.5	5.5	GenBank
376	gcccgagcgccctctgcgcgaacgcatctgagcgcgcgcgaagtggtgagcagctagtgagc		Human	435	50.0	5.8	GenBank
238	GCCGGGGCGCCCTGTGTCCCAACGCGCATGTGGCGCGCGCAAGTGTGTGCGCCCAACGGCG		Human	297	34.5	5.5	GenBank
436	ccgagcttcgcctcattcccgagcgcctaccgagcgcaagcggttgcaactcttgtctccg		Human	495	57.0	6.0	GenBank
298	CCGACTTCCTCCCTGCAATCCCGAACCGCTACCGCGCGCCACGGGCTGCACTGTGTGTCTG		Human	357	41.0	5.5	GenBank
496	gtgtgtgagcgccgcgcgcgcgaagtgcgcttggttgctctgctgaagtgcgaagcgcc		Human	555	64.0	6.5	GenBank
358	GCGCGCGCGCGCGCGCGCGCGCGCAAGGTGCGGTGCTGTGCTGTGCAAGTGCACAGCGCC		Human	417	48.0	5.5	GenBank
556	tcaccgcctcccaaccaccagctcgagctcaaggaactctggagaccgagcgccctcgccgc		Human	615	71.0	7.5	GenBank
418	TCACTGCTCTTCACACACAGCTCCGAGCTTAAGACTTGTGGGCGCGGAGCGCGCGGCGCG		Human	477	55.0	6.0	GenBank
616	agaagagcgcggaagcgccgc		Human	670	78.0	8.0	GenBank
478	AAACGGCGCGGAGCG		Human	532	62.0	6.5	GenBank

Query Match	16.2%;	Score 372.2;	DB 94;	Length 17423;
Best Local Similarity	85.9%;	Pred. NO. 1.1e-61;		
Matches 413; Conservative	0;	Mismatches 68;	Indels 0;	Gaps 0;

OY	233	agggagggtcccccacaccacctttagaaccaaaagctgctccagatcagctgcgcga	292
Db	10707	AGGCTGCAGGCCCTTCACCAATCCCTCTCTCCGACAATGTGTCCAGTACAGTGTCCCGGA	10766
OY	293	gctgaactcacaccgcgtacgtfagaccatgagcgcgctgcagcgcaagccggttacgga	352
Db	10767	GCTGACTACACAAOCGCTTCTTGACAGAGCGCCATTGGCCGACGGCAGAAGCCGGTACCBA	10826
OY	353	gctagtgtgtcccgagccaagtgcgcgccgagcgccgtctgtgtcccaagccatctgcgcg	412
Db	10827	GTTGGTGTGCTCCGGCCACAGTGGGGCCCCGGGGGCTGTGCTCCAACGGCCATGGGGGCGST	10886
OY	413	caagtgtgtgcgcacattagtgagccgcgaacttcgcgtcatccccgcgaacgttacgcgcga	472
Db	10887	CNAGTGTGTGGCCGCCGGAAGCACCGGATTTCCGCTGTCAATCCCGGATCGTACCGCCGCA	10946
OY	473	gcgcgtgtcagctgtctgtgttcccggtgtgtagaagcgccgcgcgcgaagtgcgcctggt	532
Db	10947	GCGGGTGCAGCTGTGTGGCCCCGGGGGCGCGCGCGCTCGCGCAAGTGTGCTGTGTT	11006
OY	533	ggcctctgtgcgaatgtcgaagcgctaccgcgcttccacaacccagtttgtgaacttaagactt	592
Db	11007	GCCCTGTGTGCAAGTGTCAACCGCTCACCCGTTTCCAACAACAGTGGAGCTCAAGGACTT	11066
OY	593	cggagccgaagcgctctcgcgcgcgaagaagcgccgaagccgcgcgcgcgcgcgaagcgc	652
Db	11067	CGGGCGGAGACCGGGGGCGCCGACAGAAAGGTGTGCAAAAGCGGGGCCCGGGGGGAGC	11126
OY	653	caaagccaaccagcgccgaagctgtgagagaagcctaataagagccgcgcgcctccaccac	712
Db	11127	CAAAAGCAACCGAGCGGAGACTGTGAGAACGCTACTAGACGTGAAGCCCGCTATGTGACGCC	11186
OY	713	c 713	
Db	11187	c 11187	

RESULT 14  
AC068782.2

MPCOMMENT

Sequence split into 5 fragments LOCUS AC068782 Accession AC068782

Fragment Name	Begin	End
AC068782_0	1	110000
AC068782_1	100001	210000
AC068782_2	200001	310000
AC068782_3	300001	410000
AC068782_4	400001	494838

Continuation (3 of 5) of AC068782 from base 200001 (AC068782 Mus musculus chromosome 11)

[illegible]

Consensus quality: 181484 bases at least Q40  
Consensus quality: 190159 bases at least Q30  
Consensus quality: 193745 bases at least Q20  
Insert size: 207000; agarose-fp  
Insert size: 197427; sum-of-contrigs  
Quality coverage: 3.8 in Q20 bases; agarose-fp  
Quality coverage: 4.0 in Q20 bases; sum-of-contrigs

NOTE: This is a 'working draft' sequence. It currently consists of 34 contrigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contrigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 1091: contrig of 1091 bp in length  
1092 1191: gap of 100 bp  
1192 3065: contrig of 1874 bp in length  
3066 3165: gap of 100 bp  
3166 4305: contrig of 1140 bp in length  
4306 4405: gap of 100 bp  
4406 5983: contrig of 1578 bp in length  
5984 6083: gap of 100 bp  
6084 8322: contrig of 2239 bp in length  
8323 8423: gap of 100 bp  
8423 10088: contrig of 1676 bp in length  
10089 10198: gap of 100 bp  
10199 12592: contrig of 2394 bp in length  
12593 12692: gap of 100 bp  
12693 14684: contrig of 1992 bp in length  
14685 14784: gap of 100 bp  
14785 17163: contrig of 2379 bp in length  
17164 17263: gap of 100 bp  
17264 19900: contrig of 2637 bp in length  
19901 20000: gap of 100 bp  
20001 23054: contrig of 3054 bp in length  
23055 23154: gap of 100 bp  
23155 26011: contrig of 2857 bp in length  
26012 26111: gap of 100 bp  
26112 29412: contrig of 3301 bp in length  
29413 29512: gap of 100 bp  
29513 32991: contrig of 3479 bp in length  
32992 33091: gap of 100 bp  
33092 36913: contrig of 3822 bp in length  
36914 37013: gap of 100 bp  
37014 40880: contrig of 3867 bp in length  
40881 40980: gap of 100 bp  
40981 45015: contrig of 4035 bp in length  
45016 45115: gap of 100 bp  
45116 50321: contrig of 5206 bp in length  
50322 50421: gap of 100 bp  
50422 54284: contrig of 3863 bp in length  
54285 54384: gap of 100 bp  
54385 59430: contrig of 5046 bp in length  
59431 59530: gap of 100 bp  
59531 64373: contrig of 4843 bp in length  
64374 64473: gap of 100 bp  
64474 70503: contrig of 6030 bp in length  
70504 70603: gap of 100 bp  
70604 77829: contrig of 7226 bp in length  
77830 77929: gap of 100 bp  
77930 85976: contrig of 8047 bp in length  
85977 86076: gap of 100 bp  
86077 91694: contrig of 5618 bp in length  
91695 91794: gap of 100 bp  
91795 99208: contrig of 7414 bp in length  
99209 99308: gap of 100 bp  
99309 107493: contrig of 8191 bp in length  
107500 107599: gap of 100 bp  
107600 115731: contrig of 8132 bp in length  
115732 115831: gap of 100 bp  
115832 123297: contrig of 7466 bp in length

FEATURES  
source

123298 123397: gap of 100 bp  
123398 133979: contrig of 10862 bp in length  
133980 134079: gap of 100 bp  
134080 144631: contrig of 10552 bp in length  
144632 144731: gap of 100 bp  
144732 159118: contrig of 14387 bp in length  
159119 159218: gap of 100 bp  
159219 172437: contrig of 13219 bp in length  
172438 172537: gap of 100 bp  
172538 200727: contrig of 28190 bp in length.  
Location/Qualifiers  
1. .200727  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/chromosome="11"  
/map="11"  
/clone="RP23-346P7"  
/clone\_11b="RPCI-23 Female Mouse BAC"  
1. .1091  
/note="assembly-fragment"  
1192. .3065  
/note="assembly-fragment"  
3166. .4305  
/note="assembly-fragment"  
4406. .5983  
/note="assembly-fragment"  
6084. .8322  
/note="assembly-fragment"  
8423. .10098  
/note="assembly-fragment"  
10199. .12592  
/note="assembly-fragment"  
12693. .14684  
/note="assembly-fragment"  
14785. .17163  
/note="assembly-fragment"  
17264. .19900  
/note="assembly-fragment"  
20001. .23054  
/note="assembly-fragment"  
23155. .26011  
/note="assembly-fragment"  
29413. .29512  
/note="assembly-fragment"  
29513. .32991  
/note="assembly-fragment"  
33092. .36913  
/note="assembly-fragment"  
37014. .40880  
/note="assembly-fragment"  
40981. .45015  
/note="assembly-fragment"  
45116. .50321  
/note="assembly-fragment"  
50422. .54284  
/note="assembly-fragment"  
54385. .59430  
/note="assembly-fragment"  
59531. .64373  
/note="assembly-fragment"  
64474. .70503  
/note="assembly-fragment"  
70604. .77829  
/note="assembly-fragment"  
77930. .85976  
/note="assembly-fragment"  
86077. .91694  
/note="assembly-fragment"  
91795. .99208  
/note="assembly-fragment"  
99309. .107499  
/note="assembly-fragment"



```

/note="assembly_fragment"
misc_feature      107600..115731
/note="assembly_fragment"
misc_feature      115832..123297
/note="assembly_fragment
clone_end:T7

```

Query Match 16.2%; Score 372.2; DB 62; Length 200727;  
 Best Local Similarity 85.9%; Pred. No. 7.7e-62;  
 Matches 413; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

```

OY 233 agggcgagcctccaccacccttggagaccagaagctgtccaggtacagctgcgcgca 292
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 150286 AGGCTGACGCCCTTCACGACATCCCTCTCTCCGACAGATGTCCGAGTACAGTCCGCA 150345

OY 293 gctgacttaccgctacgtgacagatggcgctgcccgaagcgaagccggtacacga 352
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
150346 GCTGCACTACACCCGCTTCTCTGACAGACGGCCCATGCGCGCAGCCGAGCGGTACACGA 150405

      353 gctggtgtgtctcgcccaagtgcgagcccgcgctgtctgccaacgcatcgccgcgca 412
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 150406 GTTGCTGTGCTCCGGCCAGTCCGCCGCCGCGCTCTGCTCCACACGCCATCGGCGCGT 150465

OY 413 caagtgtgtgacacgtatgtgacccgacttcgctgcatcccgacgcgtacgcgcgca 472
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 150466 GAAGTGTGTGGCGGCCGAAACGACCGGATTTCCGCTGCATCCGGATCGTACCGCGCGCA 150525

OY 473 ggcggtgacgctgtgtgtcccggtgtgtaggcgcgcgcgcgcgaaggtgcgcctgt 532
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 150526 GCGGGTGTGAGCTGTGTGCCCGGGGGGCGCGCGCGCTCGCGCAAGGTGCTGTGT 150585

OY 533 ggcctcgtgcaagtgcgaagcgccctaccgcgctccacaaccagtcgagctcaagactt 592
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 150586 GGCTCTGTCGCAAGTGCAGAGCGCTCACCCGCTTCACACACCACTCGAGCTCAAGGACTT 150645

OY 593 cggagccgagcgctgcgcgcgaagggcggaaagcgcgcccgcccgccgagagcgc 652
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 150646 CGGGCCGGAGACCGCGCGCGCGAGAGGGTCCCAAGCCCGCGCGCGCGGGGAGC 150705

OY 653 caagaccacacgacgcgagctgagagacgctactagagccgcgcgcgcctcccaac 712
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 150706 CAAAGCCAAACGAGCGAGCTGGAGAACGCTACTAGAGCGAGCCCGCGCTATGCAGCC 150765

OY 713 c 713
      |
150766 C 150766

```

Search completed: November 9, 2001, 00:00:48  
 Job time: 16818 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 8, 2001, 20:59:31 ; Search time 5549.44 Seconds  
(without alignments)  
906.203 Million cell updates/sec

Title: US-09-668-021-15

Perfect score: 532  
Sequence: 1 agaatgatgcacacagaatc.....accaagccagcgccgcga 532

Scoring table: IDENTITY\_NDC  
Gapop 10.0 , Gapext 1.0

Archived: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*  
10: gb\_est10:\*  
11: gb\_est11:\*  
12: gb\_est12:\*  
13: gb\_est13:\*  
14: gb\_est14:\*  
15: gb\_est15:\*  
16: gb\_est16:\*  
17: gb\_est17:\*  
18: gb\_est18:\*  
19: gb\_est19:\*  
20: gb\_est20:\*  
21: gb\_est21:\*  
22: gb\_est22:\*  
23: gb\_est23:\*  
24: gb\_est24:\*  
25: gb\_est25:\*  
26: gb\_est26:\*  
27: gb\_est27:\*  
28: gb\_est28:\*  
29: gb\_est29:\*  
30: gb\_est30:\*  
31: gb\_est31:\*  
32: gb\_est32:\*  
33: gb\_est33:\*  
34: gb\_est34:\*  
35: gb\_est35:\*  
36: gb\_est36:\*  
37: gb\_est37:\*  
38: gb\_est38:\*  
39: gb\_est39:\*  
40: gb\_est40:\*  
41: gb\_est41:\*  
42: gb\_est42:\*  
43: gb\_est43:\*  
44: gb\_est44:\*  
45: gb\_est45:\*  
46: gb\_est46:\*  
47: gb\_est47:\*  
48: gb\_est48:\*  
49: gb\_est49:\*  
50: gb\_est50:\*  
51: gb\_est51:\*  
52: gb\_est52:\*  
53: gb\_est53:\*  
54: gb\_est54:\*  
55: gb\_est55:\*  
56: gb\_est56:\*  
57: gb\_est57:\*  
58: gb\_est58:\*  
59: gb\_est59:\*  
60: gb\_est60:\*  
61: gb\_est61:\*  
62: gb\_est62:\*  
63: gb\_est63:\*  
64: gb\_est64:\*  
65: gb\_est65:\*  
66: gb\_est66:\*  
67: gb\_est67:\*  
68: gb\_est68:\*  
69: gb\_est69:\*  
70: gb\_est70:\*  
71: gb\_est71:\*  
72: gb\_est72:\*  
73: gb\_est73:\*  
74: gb\_est74:\*  
75: gb\_est75:\*  
76: gb\_est76:\*  
77: gb\_est77:\*  
78: gb\_est78:\*  
79: gb\_est79:\*  
80: gb\_est80:\*  
81: gb\_est81:\*  
82: gb\_est82:\*  
83: gb\_est83:\*  
84: gb\_est84:\*  
85: gb\_est85:\*  
86: gb\_est86:\*  
87: gb\_est87:\*  
88: gb\_est88:\*  
89: gb\_est89:\*  
90: gb\_est90:\*  
91: gb\_est91:\*  
92: gb\_est92:\*  
93: gb\_est93:\*  
94: gb\_est94:\*  
95: gb\_est95:\*  
96: gb\_est96:\*  
97: gb\_est97:\*  
98: gb\_est98:\*  
99: gb\_est99:\*  
100: gb\_est100:\*  
101: gb\_est101:\*  
102: gb\_est102:\*  
103: gb\_est103:\*  
104: gb\_est104:\*  
105: gb\_est105:\*  
106: gb\_est106:\*  
107: gb\_est107:\*  
108: gb\_est108:\*  
109: gb\_est109:\*  
110: gb\_est110:\*  
111: gb\_est111:\*  
112: gb\_est112:\*  
113: gb\_est113:\*  
114: gb\_est114:\*  
115: gb\_est115:\*  
116: gb\_est116:\*

44: em\_esthum10:\*  
45: em\_esthum11:\*  
46: em\_esthum12:\*  
47: em\_esthum13:\*  
48: em\_esthum14:\*  
49: em\_esthum15:\*  
50: em\_esthum16:\*  
51: em\_esthum17:\*  
52: em\_esthum18:\*  
53: em\_esthum19:\*  
54: em\_esthum20:\*  
55: em\_esthum21:\*  
56: em\_esthum22:\*  
57: em\_esthum23:\*  
58: em\_esthum24:\*  
59: em\_esthum25:\*  
60: em\_esthum26:\*  
61: em\_esthum27:\*  
62: em\_esthum28:\*  
63: em\_esthum29:\*  
64: em\_esthum30:\*  
65: em\_esthum31:\*  
66: em\_esthum32:\*  
67: em\_esthum33:\*  
68: em\_esthum34:\*  
69: em\_esthum35:\*  
70: em\_esthum36:\*  
71: em\_esthum37:\*  
72: em\_esthum38:\*  
73: em\_esthum39:\*  
74: em\_esthum40:\*  
75: em\_esthum41:\*  
76: em\_esthum42:\*  
77: em\_esthum43:\*  
78: em\_esthum44:\*  
79: em\_esthum45:\*  
80: em\_esthum46:\*  
81: em\_esthum47:\*  
82: em\_esthum48:\*  
83: em\_esthum49:\*  
84: em\_esthum50:\*  
85: em\_esthum51:\*  
86: em\_esthum52:\*  
87: em\_esthum53:\*  
88: em\_esthum54:\*  
89: em\_esthum55:\*  
90: em\_esthum56:\*  
91: em\_esthum57:\*  
92: em\_esthum58:\*  
93: em\_esthum59:\*  
94: em\_esthum60:\*  
95: em\_esthum61:\*  
96: em\_esthum62:\*  
97: em\_esthum63:\*  
98: em\_esthum64:\*  
99: em\_esthum65:\*  
100: em\_esthum66:\*  
101: em\_esthum67:\*  
102: em\_esthum68:\*  
103: em\_esthum69:\*  
104: em\_esthum70:\*  
105: em\_esthum71:\*  
106: em\_esthum72:\*  
107: em\_esthum73:\*  
108: em\_esthum74:\*  
109: em\_esthum75:\*  
110: em\_esthum76:\*  
111: em\_esthum77:\*  
112: em\_esthum78:\*  
113: em\_esthum79:\*  
114: em\_esthum80:\*  
115: em\_esthum81:\*  
116: em\_esthum82:\*

117: gb\_est48:\*  
 118: gb\_est49:\*  
 119: gb\_est50:\*  
 120: gb\_est51:\*  
 121: gb\_est52:\*  
 122: gb\_est53:\*  
 123: gb\_est54:\*  
 124: gb\_est55:\*  
 125: gb\_est56:\*  
 126: gb\_est57:\*  
 127: gb\_est58:\*  
 128: gb\_est59:\*  
 129: gb\_est60:\*  
 130: gb\_est61:\*  
 131: gb\_est62:\*  
 132: gb\_est63:\*  
 133: gb\_est64:\*  
 134: gb\_est65:\*  
 135: gb\_est66:\*  
 136: gb\_est67:\*  
 137: gb\_est68:\*  
 138: gb\_est69:\*  
 139: gb\_est70:\*  
 140: gb\_est71:\*  
 141: gb\_est72:\*  
 142: gb\_est73:\*  
 143: gb\_est74:\*  
 144: gb\_est75:\*  
 145: gb\_est76:\*  
 146: gb\_est77:\*  
 147: gb\_est78:\*  
 148: gb\_est79:\*  
 149: gb\_est80:\*  
 150: gb\_est81:\*  
 151: gb\_est82:\*  
 152: gb\_est83:\*  
 153: gb\_est84:\*  
 154: gb\_est85:\*  
 155: gb\_est86:\*  
 156: gb\_est87:\*  
 157: gb\_est88:\*  
 158: gb\_est89:\*  
 159: gb\_est90:\*  
 160: gb\_est91:\*  
 161: gb\_est92:\*  
 162: gb\_est93:\*  
 163: gb\_est94:\*  
 164: gb\_est95:\*  
 165: gb\_est96:\*  
 166: gb\_est97:\*  
 167: gb\_est98:\*  
 168: gb\_est99:\*  
 169: gb\_est100:\*  
 170: gb\_est101:\*  
 171: gb\_est102:\*  
 172: gb\_est103:\*  
 173: gb\_est104:\*  
 174: gb\_est105:\*  
 175: gb\_est106:\*  
 176: gb\_est107:\*  
 177: gb\_est108:\*  
 178: gb\_est109:\*

190: gb\_est110:\*  
 191: gb\_est111:\*  
 192: gb\_hlc:\*  
 193: em\_gss\_fun:\*  
 194: em\_gss\_hum1:\*  
 195: em\_gss\_hum2:\*  
 196: em\_gss\_hum3:\*  
 197: em\_gss\_hum4:\*  
 198: em\_gss\_hum5:\*  
 199: em\_gss\_hum6:\*  
 200: em\_gss\_hum7:\*  
 201: em\_gss\_hum8:\*  
 202: em\_gss\_hum9:\*  
 203: em\_gss\_inv1:\*  
 204: em\_gss\_inv2:\*  
 205: em\_gss\_inv3:\*  
 206: em\_gss\_other:\*  
 207: em\_gss\_pln1:\*  
 208: em\_gss\_pln2:\*  
 209: em\_gss\_pro:\*  
 210: em\_gss\_rod1:\*  
 211: em\_gss\_rod2:\*  
 212: em\_gss\_rod3:\*  
 213: em\_gss\_rod4:\*  
 214: em\_gss\_rod5:\*  
 215: em\_gss\_vrt1:\*  
 216: em\_gss\_vrt2:\*  
 217: em\_gss\_vrt3:\*  
 218: gb\_gss1:\*  
 219: gb\_gss2:\*  
 220: gb\_gss3:\*  
 221: gb\_gss4:\*  
 222: gb\_gss5:\*  
 223: gb\_gss6:\*  
 224: gb\_gss7:\*  
 225: gb\_gss8:\*  
 226: gb\_gss9:\*  
 227: gb\_gss10:\*  
 228: gb\_gss11:\*  
 229: gb\_gss12:\*  
 230: gb\_gss13:\*  
 231: gb\_gss14:\*  
 232: gb\_gss15:\*  
 233: gb\_gss16:\*  
 234: gb\_gss17:\*  
 235: gb\_gss18:\*  
 236: gb\_gss19:\*  
 237: gb\_gss20:\*  
 238: gb\_gss21:\*  
 239: gb\_gss22:\*  
 240: gb\_gss23:\*  
 241: gb\_gss24:\*  
 242: gb\_gss25:\*  
 243: gb\_gss26:\*  
 244: gb\_gss27:\*  
 245: gb\_gss28:\*  
 246: gb\_gss29:\*  
 247: gb\_gss30:\*  
 248: gb\_gss31:\*  
 249: gb\_gss32:\*  
 250: gb\_gss33:\*  
 251: gb\_gss34:\*  
 252: em\_gss\_inv4:\*  
 253: em\_gss\_rod6:\*  
 254: em\_gss\_rod7:\*  
 255: em\_gss\_rod8:\*  
 256: gb\_gss35:\*  
 257: gb\_gss36:\*  
 258: gb\_gss37:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 8, 2001, 20:59:27 ; Search time 5549.44 seconds  
(without alignments)  
1148.084 Million cell updates/sec

Title: US-09-668-021-13

Perfect score: 674  
Sequence: 1 gagagaccgagtgccctcctcct.....agctggagaagcgtactag 674

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Archived: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*  
10: gb\_est10:\*  
11: gb\_est11:\*  
12: gb\_est12:\*  
13: gb\_est13:\*  
14: gb\_est14:\*  
15: gb\_est15:\*  
16: gb\_est16:\*  
17: gb\_est17:\*  
18: gb\_est18:\*  
19: gb\_est19:\*  
20: gb\_est20:\*  
21: gb\_est21:\*  
22: gb\_est22:\*  
23: gb\_est23:\*  
24: gb\_est24:\*  
25: gb\_est25:\*  
26: gb\_est26:\*  
27: gb\_est27:\*  
28: gb\_est28:\*  
29: gb\_est29:\*  
30: gb\_est30:\*  
31: gb\_est31:\*  
32: gb\_est32:\*  
33: gb\_est33:\*  
34: gb\_est34:\*  
35: gb\_est35:\*  
36: gb\_est36:\*  
37: gb\_est37:\*  
38: gb\_est38:\*  
39: gb\_est39:\*  
40: gb\_est40:\*  
41: gb\_est41:\*  
42: gb\_est42:\*  
43: gb\_est43:\*

44: em\_esthum10:\*  
45: em\_esthum11:\*  
46: em\_esthum12:\*  
47: em\_esthum13:\*  
48: em\_esthum14:\*  
49: em\_esthum15:\*  
50: em\_esthum16:\*  
51: em\_esthum17:\*  
52: em\_esthum18:\*  
53: em\_esthum19:\*  
54: em\_esthum20:\*  
55: em\_esthum21:\*  
56: em\_esthum22:\*  
57: em\_esthum23:\*  
58: em\_esthum24:\*  
59: em\_esthum25:\*  
60: em\_esthum26:\*  
61: em\_esthum27:\*  
62: em\_esthum28:\*  
63: em\_esthum29:\*  
64: em\_esthum30:\*  
65: em\_esthum31:\*  
66: em\_esthum32:\*  
67: em\_esthum33:\*  
68: em\_esthum34:\*  
69: em\_esthum35:\*  
70: em\_esthum36:\*  
71: em\_esthum37:\*  
72: em\_esthum38:\*  
73: em\_esthum39:\*  
74: em\_esthum40:\*  
75: em\_esthum41:\*  
76: em\_esthum42:\*  
77: em\_esthum43:\*  
78: em\_esthum44:\*  
79: em\_esthum45:\*  
80: em\_esthum46:\*  
81: em\_esthum47:\*  
82: em\_esthum48:\*  
83: em\_esthum49:\*  
84: em\_esthum50:\*  
85: em\_esthum51:\*  
86: em\_esthum52:\*  
87: em\_esthum53:\*  
88: em\_esthum54:\*  
89: em\_esthum55:\*  
90: em\_esthum56:\*  
91: em\_esthum57:\*  
92: em\_esthum58:\*  
93: em\_esthum59:\*  
94: em\_esthum60:\*  
95: em\_esthum61:\*  
96: em\_esthum62:\*  
97: em\_esthum63:\*  
98: em\_esthum64:\*  
99: em\_esthum65:\*  
100: em\_esthum66:\*  
101: em\_esthum67:\*  
102: em\_esthum68:\*  
103: em\_esthum69:\*  
104: em\_esthum70:\*  
105: em\_esthum71:\*  
106: em\_esthum72:\*  
107: em\_esthum73:\*  
108: em\_esthum74:\*  
109: em\_esthum75:\*  
110: em\_esthum76:\*  
111: em\_esthum77:\*  
112: em\_esthum78:\*  
113: em\_esthum79:\*  
114: em\_esthum80:\*  
115: em\_esthum81:\*  
116: em\_esthum82:\*

117: gb\_est14:\*  
 118: gb\_est14:\*  
 119: gb\_est15:\*  
 120: gb\_est15:\*  
 121: gb\_est15:\*  
 122: gb\_est15:\*  
 123: gb\_est15:\*  
 124: gb\_est15:\*  
 125: gb\_est15:\*  
 126: gb\_est15:\*  
 127: gb\_est15:\*  
 128: gb\_est15:\*  
 129: gb\_est15:\*  
 130: gb\_est15:\*  
 131: gb\_est15:\*  
 132: gb\_est15:\*  
 133: gb\_est15:\*  
 134: gb\_est15:\*  
 135: gb\_est15:\*  
 136: gb\_est15:\*  
 137: gb\_est15:\*  
 138: gb\_est15:\*  
 139: gb\_est15:\*  
 140: gb\_est15:\*  
 141: gb\_est15:\*  
 142: gb\_est15:\*  
 143: gb\_est15:\*  
 144: gb\_est15:\*  
 145: gb\_est15:\*  
 146: gb\_est15:\*  
 147: gb\_est15:\*  
 148: gb\_est15:\*  
 149: gb\_est15:\*  
 150: gb\_est15:\*  
 151: gb\_est15:\*  
 152: gb\_est15:\*  
 153: gb\_est15:\*  
 154: gb\_est15:\*  
 155: gb\_est15:\*  
 156: gb\_est15:\*  
 157: gb\_est15:\*  
 158: gb\_est15:\*  
 159: gb\_est15:\*  
 160: gb\_est15:\*  
 161: gb\_est15:\*  
 162: gb\_est15:\*  
 163: gb\_est15:\*  
 164: gb\_est15:\*  
 165: gb\_est15:\*  
 166: gb\_est15:\*  
 167: gb\_est15:\*  
 168: gb\_est15:\*  
 169: gb\_est15:\*  
 170: gb\_est15:\*  
 171: gb\_est15:\*  
 172: gb\_est15:\*  
 173: gb\_est15:\*  
 174: gb\_est15:\*  
 175: gb\_est15:\*  
 176: gb\_est15:\*  
 177: gb\_est15:\*  
 178: gb\_est15:\*  
 179: gb\_est15:\*  
 180: gb\_est15:\*  
 181: gb\_est15:\*  
 182: gb\_est15:\*  
 183: gb\_est15:\*  
 184: gb\_est15:\*  
 185: gb\_est15:\*  
 186: gb\_est15:\*  
 187: gb\_est15:\*  
 188: gb\_est15:\*  
 189: gb\_est15:\*

190: gb\_est110:\*  
 191: gb\_est111:\*  
 192: gb\_hic:\*  
 193: em\_gss\_fun:\*  
 194: em\_gss\_hum1:\*  
 195: em\_gss\_hum2:\*  
 196: em\_gss\_hum3:\*  
 197: em\_gss\_hum4:\*  
 198: em\_gss\_hum5:\*  
 199: em\_gss\_hum6:\*  
 200: em\_gss\_hum7:\*  
 201: em\_gss\_hum8:\*  
 202: em\_gss\_hum9:\*  
 203: em\_gss\_inv1:\*  
 204: em\_gss\_inv2:\*  
 205: em\_gss\_inv3:\*  
 206: em\_gss\_inv4:\*  
 207: em\_gss\_inv5:\*  
 208: em\_gss\_inv6:\*  
 209: em\_gss\_inv7:\*  
 210: em\_gss\_inv8:\*  
 211: em\_gss\_inv9:\*  
 212: em\_gss\_inv10:\*  
 213: em\_gss\_inv11:\*  
 214: em\_gss\_inv12:\*  
 215: em\_gss\_inv13:\*  
 216: em\_gss\_inv14:\*  
 217: em\_gss\_inv15:\*  
 218: em\_gss\_inv16:\*  
 219: em\_gss\_inv17:\*  
 220: em\_gss\_inv18:\*  
 221: em\_gss\_inv19:\*  
 222: em\_gss\_inv20:\*  
 223: em\_gss\_inv21:\*  
 224: em\_gss\_inv22:\*  
 225: em\_gss\_inv23:\*  
 226: em\_gss\_inv24:\*  
 227: em\_gss\_inv25:\*  
 228: em\_gss\_inv26:\*  
 229: em\_gss\_inv27:\*  
 230: em\_gss\_inv28:\*  
 231: em\_gss\_inv29:\*  
 232: em\_gss\_inv30:\*  
 233: em\_gss\_inv31:\*  
 234: em\_gss\_inv32:\*  
 235: em\_gss\_inv33:\*  
 236: em\_gss\_inv34:\*  
 237: em\_gss\_inv35:\*  
 238: em\_gss\_inv36:\*  
 239: em\_gss\_inv37:\*  
 240: em\_gss\_inv38:\*  
 241: em\_gss\_inv39:\*  
 242: em\_gss\_inv40:\*  
 243: em\_gss\_inv41:\*  
 244: em\_gss\_inv42:\*  
 245: em\_gss\_inv43:\*  
 246: em\_gss\_inv44:\*  
 247: em\_gss\_inv45:\*  
 248: em\_gss\_inv46:\*  
 249: em\_gss\_inv47:\*  
 250: em\_gss\_inv48:\*  
 251: em\_gss\_inv49:\*  
 252: em\_gss\_inv50:\*  
 253: em\_gss\_inv51:\*  
 254: em\_gss\_inv52:\*  
 255: em\_gss\_inv53:\*  
 256: em\_gss\_inv54:\*  
 257: em\_gss\_inv55:\*  
 258: em\_gss\_inv56:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 8, 2001, 20:59:24 ; Search time 5549.44 Seconds  
(without alignments)  
1086.762 Million cell updates/sec

Title: US-09-668-021-11  
Perfect score: 638

Sequence: 1 atgcagccctactagcccc.....ctgagaaacctactagag 638

Scoring table: IDENTITY-NUC  
Gapop 10.0, Gapext 1.0

Archived: 10228115 segs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq Length: 0

Maximum DB seq Length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*  
10: gb\_est10:\*  
11: gb\_est11:\*  
12: gb\_est12:\*  
13: gb\_est13:\*  
14: gb\_est14:\*  
15: gb\_est15:\*  
16: gb\_est16:\*  
17: gb\_est17:\*  
18: gb\_est18:\*  
19: gb\_est19:\*  
20: gb\_est20:\*  
21: gb\_est21:\*  
22: gb\_est22:\*  
23: gb\_est23:\*  
24: gb\_est24:\*  
25: gb\_est25:\*  
26: gb\_est26:\*  
27: gb\_est27:\*  
28: gb\_est28:\*  
29: gb\_est29:\*  
30: gb\_est30:\*  
31: gb\_est31:\*  
32: gb\_est32:\*  
33: gb\_est33:\*  
34: gb\_est34:\*  
35: gb\_est35:\*  
36: gb\_est36:\*  
37: gb\_est37:\*  
38: gb\_est38:\*  
39: gb\_est39:\*  
40: gb\_est40:\*  
41: gb\_est41:\*  
42: gb\_est42:\*  
43: gb\_est43:\*

44: em\_esthum10:\*  
45: em\_esthum11:\*  
46: em\_esthum12:\*  
47: em\_esthum13:\*  
48: em\_esthum14:\*  
49: em\_esthum15:\*  
50: em\_esthum16:\*  
51: em\_esthum17:\*  
52: em\_esthum18:\*  
53: em\_esthum19:\*  
54: em\_esthum20:\*  
55: em\_esthum21:\*  
56: em\_esthum22:\*  
57: em\_esthum23:\*  
58: em\_esthum24:\*  
59: em\_esthum25:\*  
60: em\_esthum26:\*  
61: em\_esthum27:\*  
62: em\_esthum28:\*  
63: em\_esthum29:\*  
64: em\_esthum30:\*  
65: em\_esthum31:\*  
66: em\_esthum32:\*  
67: em\_esthum33:\*  
68: em\_esthum34:\*  
69: em\_esthum35:\*  
70: em\_esthum36:\*  
71: em\_esthum37:\*  
72: em\_esthum38:\*  
73: em\_esthum39:\*  
74: em\_esthum40:\*  
75: em\_esthum41:\*  
76: em\_esthum42:\*  
77: em\_esthum43:\*  
78: em\_esthum44:\*  
79: em\_esthum45:\*  
80: em\_esthum46:\*  
81: em\_esthum47:\*  
82: em\_esthum48:\*  
83: em\_esthum49:\*  
84: em\_esthum50:\*  
85: em\_esthum51:\*  
86: em\_esthum52:\*  
87: em\_esthum53:\*  
88: em\_esthum54:\*  
89: em\_esthum55:\*  
90: em\_esthum56:\*  
91: em\_esthum57:\*  
92: em\_esthum58:\*  
93: em\_esthum59:\*  
94: em\_esthum60:\*  
95: em\_esthum61:\*  
96: em\_esthum62:\*  
97: em\_esthum63:\*  
98: em\_esthum64:\*  
99: em\_esthum65:\*  
100: em\_esthum66:\*  
101: em\_esthum67:\*  
102: em\_esthum68:\*  
103: em\_esthum69:\*  
104: em\_esthum70:\*  
105: em\_esthum71:\*  
106: em\_esthum72:\*  
107: em\_esthum73:\*  
108: em\_esthum74:\*  
109: em\_esthum75:\*  
110: em\_esthum76:\*  
111: em\_esthum77:\*  
112: em\_esthum78:\*  
113: em\_esthum79:\*  
114: em\_esthum80:\*  
115: em\_esthum81:\*  
116: em\_esthum82:\*

117: gb\_est48:\*  
118: gb\_est49:\*  
119: gb\_est50:\*  
120: gb\_est51:\*  
121: gb\_est52:\*  
122: gb\_est53:\*  
123: gb\_est54:\*  
124: gb\_est55:\*  
125: gb\_est56:\*  
126: gb\_est57:\*  
127: gb\_est58:\*  
128: gb\_est59:\*  
129: gb\_est60:\*  
130: gb\_est61:\*  
131: gb\_est62:\*  
132: gb\_est63:\*  
133: gb\_est64:\*  
134: gb\_est65:\*  
135: gb\_est66:\*  
136: gb\_est67:\*  
137: gb\_est68:\*  
138: gb\_est69:\*  
139: gb\_est70:\*  
140: gb\_est71:\*  
141: gb\_est72:\*  
142: gb\_est73:\*  
143: gb\_est74:\*  
144: gb\_est75:\*  
145: gb\_est76:\*  
146: gb\_est77:\*  
147: gb\_est78:\*  
148: gb\_est79:\*  
149: gb\_est80:\*  
150: gb\_est81:\*  
151: gb\_est82:\*  
152: gb\_est83:\*  
153: gb\_est84:\*  
154: gb\_est85:\*  
155: gb\_est86:\*  
156: gb\_est87:\*  
157: gb\_est88:\*  
158: gb\_est89:\*  
159: gb\_est90:\*  
160: gb\_est91:\*  
161: gb\_est92:\*  
162: gb\_est93:\*  
163: gb\_est94:\*  
164: gb\_est95:\*  
165: gb\_est96:\*  
166: gb\_est97:\*  
167: gb\_est98:\*  
168: gb\_est99:\*  
169: gb\_est100:\*  
170: gb\_est101:\*  
171: gb\_est102:\*  
172: gb\_est103:\*  
173: gb\_est104:\*  
174: gb\_est105:\*  
175: gb\_est106:\*  
176: gb\_est107:\*  
177: gb\_est108:\*  
178: gb\_est109:\*  
179: gb\_est110:\*  
180: gb\_est111:\*  
181: gb\_est112:\*  
182: gb\_est113:\*  
183: gb\_est114:\*  
184: gb\_est115:\*  
185: gb\_est116:\*  
186: gb\_est117:\*  
187: gb\_est118:\*  
188: gb\_est119:\*  
189: gb\_est120:\*

190: gb\_est110:\*  
191: gb\_est111:\*  
192: gb\_est112:\*  
193: gb\_est113:\*  
194: gb\_est114:\*  
195: gb\_est115:\*  
196: gb\_est116:\*  
197: gb\_est117:\*  
198: gb\_est118:\*  
199: gb\_est119:\*  
200: gb\_est120:\*  
201: gb\_est121:\*  
202: gb\_est122:\*  
203: gb\_est123:\*  
204: gb\_est124:\*  
205: gb\_est125:\*  
206: gb\_est126:\*  
207: gb\_est127:\*  
208: gb\_est128:\*  
209: gb\_est129:\*  
210: gb\_est130:\*  
211: gb\_est131:\*  
212: gb\_est132:\*  
213: gb\_est133:\*  
214: gb\_est134:\*  
215: gb\_est135:\*  
216: gb\_est136:\*  
217: gb\_est137:\*  
218: gb\_est138:\*  
219: gb\_est139:\*  
220: gb\_est140:\*  
221: gb\_est141:\*  
222: gb\_est142:\*  
223: gb\_est143:\*  
224: gb\_est144:\*  
225: gb\_est145:\*  
226: gb\_est146:\*  
227: gb\_est147:\*  
228: gb\_est148:\*  
229: gb\_est149:\*  
230: gb\_est150:\*  
231: gb\_est151:\*  
232: gb\_est152:\*  
233: gb\_est153:\*  
234: gb\_est154:\*  
235: gb\_est155:\*  
236: gb\_est156:\*  
237: gb\_est157:\*  
238: gb\_est158:\*  
239: gb\_est159:\*  
240: gb\_est160:\*  
241: gb\_est161:\*  
242: gb\_est162:\*  
243: gb\_est163:\*  
244: gb\_est164:\*  
245: gb\_est165:\*  
246: gb\_est166:\*  
247: gb\_est167:\*  
248: gb\_est168:\*  
249: gb\_est169:\*  
250: gb\_est170:\*  
251: gb\_est171:\*  
252: gb\_est172:\*  
253: gb\_est173:\*  
254: gb\_est174:\*  
255: gb\_est175:\*  
256: gb\_est176:\*  
257: gb\_est177:\*  
258: gb\_est178:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 8, 2001, 20:59:19 ; Search time 5549.44 Seconds  
(without alignments)  
1093.576 Million cell updates/sec

Title: US-09-668-021-9  
Perfect score: 642  
Sequence: 1 atgcagctccactgacct.....agctgagagacgcctactag 642

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 1022815 seqs, 4726426750 residues  
Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*  
10: gb\_est10:\*  
11: gb\_est11:\*  
12: gb\_est12:\*  
13: gb\_est13:\*  
14: gb\_est14:\*  
15: gb\_est15:\*  
16: gb\_est16:\*  
17: gb\_est17:\*  
18: gb\_est18:\*  
19: gb\_est19:\*  
20: gb\_est20:\*  
21: gb\_est21:\*  
22: gb\_est22:\*  
23: gb\_est23:\*  
24: gb\_est24:\*  
25: gb\_est25:\*  
26: gb\_est26:\*  
27: gb\_est27:\*  
28: gb\_est28:\*  
29: gb\_est29:\*  
30: gb\_est30:\*  
31: gb\_est31:\*  
32: gb\_est32:\*  
33: em\_est33a:\*  
34: em\_est33b:\*  
35: em\_est33c:\*  
36: em\_est33d:\*  
37: em\_est33e:\*  
38: em\_est33f:\*  
39: em\_est33g:\*  
40: em\_est33h:\*  
41: em\_est33i:\*  
42: em\_est33j:\*  
43: em\_est33k:\*

X

44: em\_esthum10:\*  
45: em\_esthum11:\*  
46: em\_esthum12:\*  
47: em\_esthum13:\*  
48: em\_esthum14:\*  
49: em\_esthum15:\*  
50: em\_esthum16:\*  
51: em\_esthum17:\*  
52: em\_esthum18:\*  
53: em\_esthum19:\*  
54: em\_esthum20:\*  
55: em\_esthum21:\*  
56: em\_esthum22:\*  
57: em\_esthum23:\*  
58: em\_esthum24:\*  
59: em\_esthum25:\*  
60: em\_esthum26:\*  
61: em\_esthum27:\*  
62: em\_esthum28:\*  
63: em\_esthum29:\*  
64: em\_esthum30:\*  
65: em\_esthum31:\*  
66: em\_esthum32:\*  
67: em\_esthum33:\*  
68: em\_esthum34:\*  
69: em\_esthum35:\*  
70: em\_esthum36:\*  
71: em\_esthum37:\*  
72: em\_esthum38:\*  
73: em\_esthum39:\*  
74: em\_esthum40:\*  
75: em\_esthum41:\*  
76: em\_esthum42:\*  
77: em\_esthum43:\*  
78: em\_esthum44:\*  
79: em\_esthum45:\*  
80: em\_esthum46:\*  
81: em\_esthum47:\*  
82: em\_esthum48:\*  
83: em\_esthum49:\*  
84: em\_esthum50:\*  
85: em\_esthum51:\*  
86: em\_esthum52:\*  
87: em\_esthum53:\*  
88: em\_esthum54:\*  
89: em\_esthum55:\*  
90: em\_esthum56:\*  
91: em\_esthum57:\*  
92: em\_esthum58:\*  
93: em\_esthum59:\*  
94: em\_esthum60:\*  
95: em\_esthum61:\*  
96: em\_esthum62:\*  
97: em\_esthum63:\*  
98: em\_esthum64:\*  
99: em\_esthum65:\*  
100: em\_esthum66:\*  
101: em\_esthum67:\*  
102: em\_esthum68:\*  
103: em\_esthum69:\*  
104: em\_esthum70:\*  
105: em\_esthum71:\*  
106: em\_esthum72:\*  
107: em\_esthum73:\*  
108: em\_esthum74:\*  
109: em\_esthum75:\*  
110: em\_esthum76:\*  
111: em\_esthum77:\*  
112: em\_esthum78:\*  
113: em\_esthum79:\*  
114: em\_esthum80:\*  
115: em\_esthum81:\*  
116: em\_esthum82:\*

117: gb\_est48:\*  
118: gb\_est49:\*  
119: gb\_est50:\*  
120: gb\_est51:\*  
121: gb\_est52:\*  
122: gb\_est53:\*  
123: gb\_est54:\*  
124: gb\_est55:\*  
125: gb\_est56:\*  
126: gb\_est57:\*  
127: gb\_est58:\*  
128: gb\_est59:\*  
129: gb\_est60:\*  
130: gb\_est61:\*  
131: gb\_est62:\*  
132: gb\_est63:\*  
133: gb\_est64:\*  
134: gb\_est65:\*  
135: gb\_est66:\*  
136: gb\_est67:\*  
137: gb\_est68:\*  
138: gb\_est69:\*  
139: gb\_est70:\*  
140: gb\_est71:\*  
141: gb\_est72:\*  
142: gb\_est73:\*  
143: gb\_est74:\*  
144: gb\_est75:\*  
145: gb\_est76:\*  
146: gb\_est77:\*  
147: gb\_est78:\*  
148: gb\_est79:\*  
149: gb\_est80:\*  
150: gb\_est81:\*  
151: gb\_est82:\*  
152: gb\_est83:\*  
153: gb\_est84:\*  
154: gb\_est85:\*  
155: gb\_est86:\*  
156: gb\_est87:\*  
157: gb\_est88:\*  
158: gb\_est89:\*  
159: gb\_est90:\*  
160: gb\_est91:\*  
161: gb\_est92:\*  
162: gb\_est93:\*  
163: gb\_est94:\*  
164: gb\_est95:\*  
165: gb\_est96:\*  
166: gb\_est97:\*  
167: gb\_est98:\*  
168: gb\_est99:\*  
169: gb\_est100:\*  
170: gb\_est101:\*  
171: gb\_est102:\*  
172: gb\_est103:\*  
173: gb\_est104:\*  
174: gb\_est105:\*  
175: gb\_est106:\*  
176: gb\_est107:\*  
177: gb\_est108:\*  
178: gb\_est109:\*

190: gb\_est110:\*  
191: gb\_est111:\*  
192: gb\_hrc:\*  
193: em\_gss\_fun:\*  
194: em\_gss\_hum1:\*  
195: em\_gss\_hum2:\*  
196: em\_gss\_hum3:\*  
197: em\_gss\_hum4:\*  
198: em\_gss\_hum5:\*  
199: em\_gss\_hum6:\*  
200: em\_gss\_hum7:\*  
201: em\_gss\_hum8:\*  
202: em\_gss\_hum9:\*  
203: em\_gss\_inv1:\*  
204: em\_gss\_inv2:\*  
205: em\_gss\_inv3:\*  
206: em\_gss\_inv4:\*  
207: em\_gss\_inv5:\*  
208: em\_gss\_inv6:\*  
209: em\_gss\_inv7:\*  
210: em\_gss\_inv8:\*  
211: em\_gss\_inv9:\*  
212: em\_gss\_inv10:\*  
213: em\_gss\_inv11:\*  
214: em\_gss\_inv12:\*  
215: em\_gss\_inv13:\*  
216: em\_gss\_inv14:\*  
217: em\_gss\_inv15:\*  
218: em\_gss\_inv16:\*  
219: em\_gss\_inv17:\*  
220: em\_gss\_inv18:\*  
221: em\_gss\_inv19:\*  
222: em\_gss\_inv20:\*  
223: em\_gss\_inv21:\*  
224: em\_gss\_inv22:\*  
225: em\_gss\_inv23:\*  
226: em\_gss\_inv24:\*  
227: em\_gss\_inv25:\*  
228: em\_gss\_inv26:\*  
229: em\_gss\_inv27:\*  
230: em\_gss\_inv28:\*  
231: em\_gss\_inv29:\*  
232: em\_gss\_inv30:\*  
233: em\_gss\_inv31:\*  
234: em\_gss\_inv32:\*  
235: em\_gss\_inv33:\*  
236: em\_gss\_inv34:\*  
237: em\_gss\_inv35:\*  
238: em\_gss\_inv36:\*  
239: em\_gss\_inv37:\*  
240: em\_gss\_inv38:\*  
241: em\_gss\_inv39:\*  
242: em\_gss\_inv40:\*  
243: em\_gss\_inv41:\*  
244: em\_gss\_inv42:\*  
245: em\_gss\_inv43:\*  
246: em\_gss\_inv44:\*  
247: em\_gss\_inv45:\*  
248: em\_gss\_inv46:\*  
249: em\_gss\_inv47:\*  
250: em\_gss\_inv48:\*  
251: em\_gss\_inv49:\*  
252: em\_gss\_inv50:\*  
253: em\_gss\_inv51:\*  
254: em\_gss\_inv52:\*  
255: em\_gss\_inv53:\*  
256: em\_gss\_inv54:\*  
257: em\_gss\_inv55:\*  
258: em\_gss\_inv56:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 9, 2001, 00:10:40 ; Search time 9526.57 Seconds  
(without alignments)  
863.778 Million cell updates/sec

Title: US-09-668-021-15  
532

Sequence: 1 agaatgatgcacagaatc.....accgaagccagcgagccga 532

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:\*

- 1: gb\_ba1:\*
- 2: gb\_ba2:\*
- 3: gb\_ba3:\*
- 4: gb\_in1:\*
- 5: gb\_in2:\*
- 6: gb\_in3:\*
- 7: gb\_ov:\*
- 8: gb\_ov:\*
- 9: gb\_pat1:\*
- 10: gb\_pat2:\*
- 11: gb\_ph:\*
- 12: gb\_p11:\*
- 13: gb\_p12:\*
- 14: gb\_p13:\*
- 15: gb\_p14:\*
- 16: gb\_ba1:\*
- 17: gb\_ba2:\*
- 18: em\_fun:\*
- 19: em\_hrgo\_hum:\*
- 20: em\_hrgo\_inv:\*
- 21: em\_hrgo\_rtd:\*
- 22: em\_hrg\_hum1:\*
- 23: em\_hrg\_hum2:\*
- 24: em\_hrg\_hum3:\*
- 25: em\_hrg\_hum4:\*
- 26: em\_hrg\_hum5:\*
- 27: em\_hrg\_hum6:\*
- 28: em\_hrg\_hum7:\*
- 29: em\_hrg\_hum8:\*
- 30: em\_hrg\_inv1:\*
- 31: em\_hrg\_inv2:\*
- 32: em\_hrg\_other:\*
- 33: em\_hrg\_rtd:\*
- 34: em\_hum1:\*
- 35: em\_hum2:\*
- 36: em\_hum3:\*
- 37: em\_hum4:\*
- 38: em\_hum5:\*
- 39: em\_hum6:\*
- 40: em\_hum7:\*
- 41: em\_in:\*
- 42: em\_ov:\*
- 43: em\_or:\*

44: em\_ov:\*

45: em\_pat:\*

46: em\_ph:\*

47: em\_p1:\*

48: em\_ro:\*

49: em\_sts:\*

50: em\_sy:\*

51: em\_un:\*

52: em\_v1:\*

53: gb\_sts1:\*

54: gb\_sts2:\*

55: gb\_sts3:\*

56: gb\_sy:\*

57: gb\_un:\*

58: gb\_v11:\*

59: gb\_v12:\*

60: gb\_hrg1:\*

61: gb\_hrg2:\*

62: gb\_hrg3:\*

63: gb\_hrg4:\*

64: gb\_hrg5:\*

65: gb\_hrg6:\*

66: gb\_hrg7:\*

67: gb\_hrg8:\*

68: gb\_hrg9:\*

69: gb\_hrg10:\*

70: gb\_hrg11:\*

71: gb\_hrg12:\*

72: gb\_hrg13:\*

73: gb\_hrg14:\*

74: gb\_hrg15:\*

75: gb\_hrg16:\*

76: gb\_hrg17:\*

77: gb\_hrg18:\*

78: gb\_hrg19:\*

79: gb\_hrg20:\*

80: gb\_hrg21:\*

81: gb\_hrg22:\*

82: gb\_hrg23:\*

83: gb\_hrg24:\*

84: gb\_hrg25:\*

85: gb\_p11:\*

86: gb\_p12:\*

87: gb\_p13:\*

88: gb\_p14:\*

89: gb\_p15:\*

90: gb\_p16:\*

91: gb\_p17:\*

92: gb\_p18:\*

93: gb\_p19:\*

94: gb\_p101:\*

95: gb\_p102:\*

96: gb\_p104:\*

97: gb\_p110:\*

98: em\_ba3:\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query length	DB ID	Description
1	532	100.0	532	7 AF326738	AF326738 Bos tauru
2	467.8	87.9	2236	89 AF331844	AF331844 Homo sapi
3	467.8	87.9	2323	89 AF326739	AF326739 Homo sapi
4	467.8	87.9	2329	9 AX056687	AX056687 Sequence
5	461.4	86.7	642	89 AF326742	AF326742 Cercopit
6	424.6	79.8	674	94 AF326741	AF326741 Rattus no
7	414.2	77.9	638	94 AF326740	AF326740 Mus muscu
8	375.2	70.5	21501	89 AF326736	AF326736 Homo sapi

[illegible]

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 9, 2001, 00:07:17 ; Search time 9526.57 Seconds

(without alignments)  
1094.335 Million cell updates/sec

Title: US-09-668-021-13

Perfect score: 674

Sequence: 1 gagagccgagtcctctct.....agctgagagcgcctactag 674

Scoring table: IDENTITY\_NDC  
Gapop 10.0 , Gapext 1.0

Archived: 1344157 segs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:\*

1: gb\_ba1:\*  
2: gb\_ba2:\*  
3: gb\_ba3:\*  
4: gb\_in1:\*  
5: gb\_in2:\*  
6: gb\_in3:\*  
7: gb\_om:\*  
8: gb\_ov:\*  
9: gb\_pat1:\*  
10: gb\_pat2:\*  
11: gb\_ph:\*  
12: gb\_pi1:\*  
13: gb\_pi2:\*  
14: gb\_pi3:\*  
15: gb\_pi4:\*  
16: em\_ba1:\*  
17: em\_ba2:\*  
18: em\_fun:\*  
19: em\_hcgo\_hum:\*  
20: em\_hcgo\_inv:\*  
21: em\_hcgo\_rtd:\*  
22: em\_hcgo\_hum1:\*  
23: em\_hcgo\_hum2:\*  
24: em\_hcgo\_hum3:\*  
25: em\_hcgo\_hum4:\*  
26: em\_hcgo\_hum5:\*  
27: em\_hcgo\_hum6:\*  
28: em\_hcgo\_hum7:\*  
29: em\_hcgo\_hum8:\*  
30: em\_hcgo\_inv1:\*  
31: em\_hcgo\_inv2:\*  
32: em\_hcgo\_other:\*  
33: em\_hcgo\_rtd:\*  
34: em\_hum1:\*  
35: em\_hum2:\*  
36: em\_hum3:\*  
37: em\_hum4:\*  
38: em\_hum5:\*  
39: em\_hum6:\*  
40: em\_hum7:\*  
41: em\_in:\*  
42: em\_om:\*  
43: em\_or:\*

44: em\_ov:\*  
45: em\_pat:\*  
46: em\_ph:\*  
47: em\_pi:\*  
48: em\_pi1:\*  
49: em\_pi2:\*  
50: em\_pi3:\*  
51: em\_pi4:\*  
52: em\_pi5:\*  
53: em\_pi6:\*  
54: em\_pi7:\*  
55: em\_pi8:\*  
56: em\_pi9:\*  
57: em\_pi10:\*  
58: em\_pi11:\*  
59: em\_pi12:\*  
60: em\_pi13:\*  
61: em\_pi14:\*  
62: em\_pi15:\*  
63: em\_pi16:\*  
64: em\_pi17:\*  
65: em\_pi18:\*  
66: em\_pi19:\*  
67: em\_pi20:\*  
68: em\_pi21:\*  
69: em\_pi22:\*  
70: em\_pi23:\*  
71: em\_pi24:\*  
72: em\_pi25:\*  
73: em\_pi26:\*  
74: em\_pi27:\*  
75: em\_pi28:\*  
76: em\_pi29:\*  
77: em\_pi30:\*  
78: em\_pi31:\*  
79: em\_pi32:\*  
80: em\_pi33:\*  
81: em\_pi34:\*  
82: em\_pi35:\*  
83: em\_pi36:\*  
84: em\_pi37:\*  
85: em\_pi38:\*  
86: em\_pi39:\*  
87: em\_pi40:\*  
88: em\_pi41:\*  
89: em\_pi42:\*  
90: em\_pi43:\*  
91: em\_pi44:\*  
92: em\_pi45:\*  
93: em\_pi46:\*  
94: em\_pi47:\*  
95: em\_pi48:\*  
96: em\_pi49:\*  
97: em\_pi50:\*  
98: em\_pi51:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	674	100.0	674	94	AF326741 Rattus no
2	556	82.5	638	94	AF326740 Mus muscu
3	536.2	79.6	2296	89	AF331844 Homo sapi
4	536.2	79.6	2293	89	AF326739 Homo sapi
5	536.2	79.6	2329	9	AX056687 Sequence
6	530	78.6	642	89	AF326742 Cercopit
7	424.6	63.0	532	7	AF326738 Bos tauru
8	407.4	60.4	17423	94	AF326737 Mus muscu

9	407.4	60.4	110000	73	AC0668782.2	Continuation (3 conts)
10	407.4	60.4	200727	62	AC012296	AC012296 Mus musculus
11	407.4	60.4	211533	73	AC0668807	AC0668807 Mus musculus
12	372.6	55.3	21501	89	AF326736	AF326736 Homo sapiens
13	372.6	55.3	94752	85	AC003098	AC003098 Homo sapiens
14	372.6	55.3	151780	72	AC055813	AC055813 Homo sapiens
15	288.2	42.8	177744	75	AC073954	AC073954 Homo sapiens
16	167.4	24.8	51575	68	AC023810	AC023810 Mus musculus
17	64	9.5	900	9	AR017832	AR017832 Sequence
18	62	9.2	150951	75	AC079155	AC079155 Homo sapiens
19	60.4	9.0	1500	93	BSM860018	AL050024 Homo sapiens
20	56.2	8.3	412	10	AX071502	AX071502 Sequence
21	54.2	8.0	58733	2	AF217189	AF217189 Sorangium
22	54	8.0	34611	3	SC6B3A	AL356632 Streptomyces
23	54	8.0	37856	2	AX024319	AX024319 Sequence
24	54	8.0	37856	2	AX024312	AX024312 Sequence
25	52.8	7.8	1153	53	CMS07BVP	AL438395 T3 end of
26	52.2	7.7	208978	73	AC067931	AC067931 Homo sapiens
27	51.8	7.7	33095	3	SC3C8	AL023861 Streptomyces
28	50.6	7.5	1898	3	SCRXABGN	KX2105 Streptomyces
29	50.6	7.5	41625	3	SC424	AL049626 Streptomyces
30	50.4	7.5	567	53	CMS07B91	AL429856 clone BAO
31	50.2	7.4	204224	77	AC008781	AC008781 Mus musculus
32	50	7.4	2963	4	AB044709	AB044709 Gryllus b
33	49.8	7.4	2164	3	SC2C3	AB044709 Gryllus b
34	49.4	7.3	170020	13	AP003256	AP003256 Oryza sativa
35	49.2	7.3	666	93	HSNTCH13	AF055893 Homo sapiens
36	49.2	7.3	8091	9	A69561	A69561 Sequence 1
37	49.2	7.3	8091	97	HSU97669	U97669 Homo sapiens
38	49.2	7.3	14053	3	SC2134	AL445403 Streptomyces
39	49.2	7.3	41150	85	AC004067	AC004067 Homo sapiens
40	49.2	7.3	252859	68	AC024907	AC024907 Homo sapiens
41	49	7.3	963	3	SCRXRBGN	X92104 S. coelicolor
42	49	7.3	1153	53	CMS07BVP	AL438395 T3 end of
43	48	7.3	4026	3	SC0007313	AJ007713 Streptomyces
44	48	7.1	367	73	AC042856	AL42856 clone BAO
45	48	7.1	52884	73	CMS068018	AC068018 Homo sapiens

## ALIGNMENTS

	Continuation (3 of
AC012296	Mus muscu
AC068807	Mus muscu
AJ326736	Homo sapi
AC030398	Homo sapi
AC055813	Homo sapi
AC073954	Homo sapi
AC023810	Mus muscu
AB017832	Sequence
AC079155	Homo sapi
AL050024	Homo sapi
AX071502	Sequence
AF217189	Soraria lum
AL356832	Streptomy
AX024319	Sequence
AX024212	Sequence
AL438395	T3 end of
AC067931	Homo sapi
AL023861	Streptomy
X92105	Streptomyce
AL049826	Streptomy
AL429856	clone BAO
AC087891	Mus muscu
AB044709	Gryllus b
AL391157	Streptomy
AP003256	Oryza sat
AF055893	Homo sapi
A69561	Sequence 1
U97669	Homo saplen
AL445403	Streptomy
AC004663	Homo sapi
AC024907	Homo sapi
X92104	S. coelicolo
AL438395	T3 end of
AU079713	Streptomy
AL429856	clone BAO
AC068018	Homo sapi

  

5'UTR	/organism="Rattus norvegicus"
CDS	/strain="Sprague-Dawley"
	/db_xref="taxon:10116"
	1..32
	33..674
	/codon_start=1
	/product="sclerostin"
	/protein_id="AK13456.1"
	/db_xref="GI:13161026"
	/translation="MQLSLAPKICLILIAAFVAVESQGMQAFKNMDATEIIPGIREXP
	EPPDELNNOTNRAENGSRPHNRHRYDQSEVSCRRLHTREYTPGGPSARPYE
	LVSQGCGPAPRLPAITGRKKMKRNQRPFRCTIPRTAQRVOLLCPEGANPRSKRV
	LVASCKKIKTLTRFHNSLEKDPEPTARFQKKPRPARAKMANQAELENAT"

  

BASE COUNT	128 a	240 c	208 g	98 t
ORIGIN				
Query Match	100.0%;	Score 674;	DB 94;	Length 674;
Best Local Similarity	100.0%;	Pred. No. 9,7e-112;		
Matches 674;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

  

QY	1	gaggaaccagatgccttcctcctcctcttgagcaacatgatcacgtctctaataagccctgcttg	60
Db	1	GAGGACCGAGTCCCTCCTCCTCCTTG6GCACAATGCAAGCTCATGAAGCCCTTGCTTG	60
QY	61	cctgctgccttgttatatcacgcttgcgttttcgtggaggacccagggttggaagcttica	120
Db	61	CCGCGCTGCTTTATCATCACGCTTCGTGTTGTGGAGACCAGGGGTGGCANAGCTTCA	120
QY	121	agaatgatgccacagaaatcatcccggaactcagagagtataccagaagctcctcaggaac	180
Db	121	AGAAATGATGCCACAGAATTCATCCC6GGACTCAGAGAGTACCACAGCCTCCTCAGGAAC	180
QY	181	ttagaagacaacagagccaatgaaccggagccagaggaacagagagcagacccccacacatcctt	240
Db	181	TGAGAGACACACGACCACTGAACCGGGGCTCGAGACGAGGCAACCCCCCCACATCCTT	240

FEATURES	source	1
AF326741	674 bp	MRNA
LOCUS		28-FEB-2001
DEFINITION		Rattus norvegicus sclerostin mRNA, complete cds.
ACCESSION		AF326741
VERSION		AF326741.1
KEYWORDS		GI:13161025
SOURCE		
ORGANISM		Norway rat.
		Rattus norvegicus
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
		Rattus.
REFERENCE		1 (bases 1 to 674)
AUTHORS		Brunkow,M.E., Gardner,J.C., Van Ness,J., Paepert,B.W.,
		Kovachovich,B.R., Proll,S., Skonler,J.E., Zhao,L., Sabo,P.J.,
		Fu,Y.H., Alish,R.S., Gillett,L., Colbert,T., Tacconi,P., Galas,D.
		Hamerma,H., Beighton,P., and Mulligan,J.T.
		Bone Dysplasia Sclerosteosis Results from Loss of the SOST Gene
		Product, a Novel Cystine Knot-Containing Protein
		Am. J. Hum. Genet. 68 (3), 577-589 (2001)
JOURNAL		11179006
PUBMED		2 (bases 1 to 674)
REFERENCE		Brunkow,M.E., Gardner,J.C., Van Ness,J., Paepert,B.W.,
AUTHORS		Kovachovich,B.R., Proll,S., Skonler,J.E., Zhao,L., Sabo,P.J.,
		Fu,Y.H., Alish,R.S., Gillett,L., Colbert,T., Tacconi,P., Galas,D.
		Hamerma,H., Beighton,P., and Mulligan,J.T.
		Direct Submission
		Submitted (07-DEC-2000) Genomics, Cellegch Chiroscience Inc., 1631
		220th St. SE, Bothell, WA 98021, USA
		Location/Qualifiers
		1..674

[illegible]

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 9, 2001, 00:04:04 : Search time 9526.57 Seconds  
(without alignments)  
1035.884 Million cell updates/sec

Title: US-09-668-021-11

Perfect score: 638  
Sequence: 1 atcgagccctactagcccc.....ctggagaacgcctactagag 638

Scoring table: IDENTIFY\_NUC  
Gapop 10.0 , Gapext 1.0

arched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:\*  
1: gb\_ba1:\*  
2: gb\_ba2:\*  
3: gb\_ba3:\*  
4: gb\_in1:\*  
5: gb\_in2:\*  
6: gb\_in3:\*  
7: gb\_om:\*  
8: gb\_ov:\*  
9: gb\_pat1:\*  
10: gb\_pat2:\*  
11: gb\_ph:\*  
12: gb\_pl1:\*  
13: gb\_pl2:\*  
14: gb\_pl3:\*  
15: gb\_pl4:\*  
16: em\_ba1:\*  
17: em\_ba2:\*  
18: em\_fun:\*  
19: em\_higo\_hum:\*  
20: em\_higo\_inv:\*  
21: em\_higo\_rod:\*  
22: em\_hig\_hum1:\*  
23: em\_hig\_hum2:\*  
24: em\_hig\_hum3:\*  
25: em\_hig\_hum4:\*  
26: em\_hig\_hum5:\*  
27: em\_hig\_hum6:\*  
28: em\_hig\_hum7:\*  
29: em\_hig\_hum8:\*  
30: em\_hig\_hum1:\*  
31: em\_hig\_inv2:\*  
32: em\_hig\_other:\*  
33: em\_hig\_rod:\*  
34: em\_hum1:\*  
35: em\_hum2:\*  
36: em\_hum3:\*  
37: em\_hum4:\*  
38: em\_hum5:\*  
39: em\_hum6:\*  
40: em\_hum7:\*  
41: em\_in:\*  
42: em\_om:\*  
43: em\_or:\*

44: em\_ov:\*  
45: em\_pat:\*  
46: em\_ph:\*  
47: em\_pl:\*  
48: em\_rod:\*  
49: em\_sts:\*  
50: em\_sy:\*  
51: em\_un:\*  
52: em\_v1:\*  
53: gb\_sts1:\*  
54: gb\_sts2:\*  
55: gb\_sts3:\*  
56: gb\_sy:\*  
57: gb\_un:\*  
58: gb\_v11:\*  
59: gb\_v12:\*  
60: gb\_hig1:\*  
61: gb\_hig2:\*  
62: gb\_hig3:\*  
63: gb\_hig4:\*  
64: gb\_hig5:\*  
65: gb\_hig6:\*  
66: gb\_hig7:\*  
67: gb\_hig8:\*  
68: gb\_hig9:\*  
69: gb\_hig10:\*  
70: gb\_hig11:\*  
71: gb\_hig12:\*  
72: gb\_hig13:\*  
73: gb\_hig14:\*  
74: gb\_hig15:\*  
75: gb\_hig16:\*  
76: gb\_hig17:\*  
77: gb\_hig18:\*  
78: gb\_hig19:\*  
79: gb\_hig20:\*  
80: gb\_hig21:\*  
81: gb\_hig22:\*  
82: gb\_hig23:\*  
83: gb\_hig24:\*  
84: gb\_hig25:\*  
85: gb\_pr1:\*  
86: gb\_pr2:\*  
87: gb\_pr3:\*  
88: gb\_pr4:\*  
89: gb\_pr5:\*  
90: gb\_pr6:\*  
91: gb\_pr7:\*  
92: gb\_pr8:\*  
93: gb\_pr9:\*  
94: gb\_pr10:\*  
95: gb\_pr11:\*  
96: gb\_pr12:\*  
97: gb\_pr13:\*  
98: em\_ba3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	638	100.0	638	94	AF326740 Mus muscu
2	556	87.1	674	94	AF326741 Rattus no
3	503.2	78.9	642	89	AF326742 Cercopith
4	502	78.7	2296	89	AF331844 Homo sapi
5	502	78.7	2323	89	AF326739
6	502	78.7	2329	9	AX056687 Sequence
7	427	66.9	17423	94	AF326737 Mus muscu
8	427	66.9	110000	73	AC068782_2 Continuation (3 of

Query Match	100.0% ; Score 638 ; DB 94 ; Length 638 ;
Best Local Similarity	100.0% ; Pred. NO. 6, 1e-101 ;
Matches 638 ; Conservative	0 ; Mismatches 0 ; Indels 0 ; Gaps 0 ;
QY 1 atgcagccctcaactagaccgcgtgcctcactctgcctacttgcacgctgcctctgtgct 60	/strain="1295v"/j /db_xref="taxon:10090" /chromosome="11" 1. 636 /codon_start=1 /product="sclerostin" /protein_id="AAK13455.1" /db_xref="GI:13161023" /translation="MPSLAPCLICLLVHAACFAVEGGOMAFRNDATVEVPIGIEEP EPPEENNQTMRAENGSRPPHPHYDAKGSVESECLHLYTPTLDGCRSAKPYTELY CSGGQRLPLNALIGVAKMPPNGPDFICTRIRARQVQLCPGAAIPRSKRVLY ASCKKRLFRHNSSELEKDPETAPQGRPRPGARAKANAELENA"
Db 1 ATGCAGCCCTCACTAGACCCCGCTGCCTCAWCTGCCTACTTGTCAACCTCTCTTGCTGTCT 60	
QY 61 gtggaggggccaaggggtggcaagcctctggagatgatatccacaaggtcattcccaaggtct 120	
Db 61 GTGGAGGGGCCAGGGGGTGGCAAGCCTTTCAGGATATGCCACAAAGGTCATCCAGGGCTT 120	
QY 121 gggaggttaccccgagagctctctctcttgaagaacaacagacacgtgaacccggcgagaaatgaa 180	
Db 121 GGAGAGTACCCCGAGGCTCTCTCTGTGAGAAACACGACCATGAAACCGGGCGGGAATGGA 180	
QY 181 ggcagagcctccacacatccctatgacgccaaggtgtctcgaagtatagagctggccggag 240	
Db 181 GGCAGACCTCCACCATCCCTATGACGCCAAGAGTGTGTCTCCAGTACACTTCCGCCAG 240	
QY 241 ctgcactaacaccgctctctctgacagacggccaatgcccagcgcgcgaacccggttcacggag 300	



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 9, 2001, 00:00:48 ; Search time 9526.57 Seconds

(without alignments)  
1042.379 Million cell updates/sec

Title: US-09-668-021-9

Perfect score: 642

Sequence: 1 atgcagctccactgagcct.....agctgagagaagcactactag 642

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
GenBmb1.*
1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_om:*
8: gb_om:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_pl1:*
13: gb_pl2:*
14: gb_pl3:*
15: gb_pl4:*
16: em_ba1:*
17: em_ba2:*
18: em_fun:*
19: em_htgo_hum:*
20: em_htgo_inv:*
21: em_htgo_rod:*
22: em_htg_hum1:*
23: em_htg_hum2:*
24: em_htg_hum3:*
25: em_htg_hum4:*
26: em_htg_hum5:*
27: em_htg_hum6:*
28: em_htg_hum7:*
29: em_htg_hum8:*
30: em_htg_inv1:*
31: em_htg_inv2:*
32: em_htg_other:*
33: em_htg_rod:*
34: em_hum1:*
35: em_hum2:*
36: em_hum3:*
37: em_hum4:*
38: em_hum5:*
39: em_hum6:*
40: em_hum7:*
41: em_in:*
42: em_om:*
43: em_or:*
```

```
44: em_ov:*
45: em_pat:*
46: em_ph:*
47: em_pl:*
48: em_rod:*
49: em_sts:*
50: em_sy:*
51: em_un:*
52: em_v1:*
53: gb_sts1:*
54: gb_sts2:*
55: gb_sts3:*
56: gb_sy:*
57: gb_un:*
58: gb_v1:*
59: gb_v12:*
60: gb_hcg1:*
61: gb_hcg2:*
62: gb_hcg3:*
63: gb_hcg4:*
64: gb_hcg5:*
65: gb_hcg6:*
66: gb_hcg7:*
67: gb_hcg8:*
68: gb_hcg9:*
69: gb_hcg10:*
70: gb_hcg11:*
71: gb_hcg12:*
72: gb_hcg13:*
73: gb_hcg14:*
74: gb_hcg15:*
75: gb_hcg16:*
76: gb_hcg17:*
77: gb_hcg18:*
78: gb_hcg19:*
79: gb_hcg20:*
80: gb_hcg21:*
81: gb_hcg22:*
82: gb_hcg23:*
83: gb_hcg24:*
84: gb_hcg25:*
85: gb_pr1:*
86: gb_pr2:*
87: gb_pr3:*
88: gb_pr4:*
89: gb_pr5:*
90: gb_pr6:*
91: gb_pr7:*
92: gb_pr8:*
93: gb_pr9:*
94: gb_rol:*
95: gb_rol:*
96: gb_rol:*
97: gb_rol:*
98: em_ba3:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	642	100.0	642	89	AF326742	AF326742 Cercopit
2	613.2	95.5	2296	89	AF331844	AF331844 Homo sapi
3	613.2	95.5	2329	89	AF326739	AF326739 Homo sapi
4	613.2	95.5	2329	9	AX056687	AX056687 Sequence
5	530	82.6	674	94	AF326741	AF326741 Rattus no
6	503.2	78.4	638	94	AF326740	AF326740 Mus muscu
7	461.4	71.9	532	7	AF326738	AF326738 Bos tauru
8	413.4	64.4	21501	89	AF326736	AF326736 Homo sapi

